

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Computer: Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 15:52:26 : Search time 35 seconds
(without alignments)
2566 788 Million cell updates/sec

Title: US-09-866-379a-10
Perfect score: 2279
Sequence: 1 MKALLPFLSLIPLTPQSA.....GFTQIVNEARIPACSLRSH 436

Scoring table: BLUSMB2

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 20647115 residues 671580

Total number of hits satisfying chosen parameters.

Minimum DB seq length: 0

Maximum DB seq length: 206000000

Post-processing: Minimum Match 6%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_oranalle.*

9: sp_plaque.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_virus.*

16: sp_bacteriap.*

17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 2176 | 95.5 | 432 | Q9RFE3 | Q9RFE3 escherichia |
| 2 | 2176 | 95.5 | 432 | Q9RFE3 | Q9RFE3 escherichia |
| 3 | 2176 | 95.5 | 432 | Q9RFE3 | Q9RFE3 escherichia |
| 4 | 2176 | 95.5 | 432 | Q9RFE3 | Q9RFE3 escherichia |
| 5 | 2176 | 95.5 | 432 | Q9RFE3 | Q9RFE3 escherichia |
| 6 | 2172 | 95.3 | 432 | Q9RFE3 | Q9RFE3 escherichia |
| 7 | 2144 | 94.1 | 434 | Q9RFE3 | Q9RFE3 escherichia |
| 8 | 959.5 | 42.1 | 441 | Q9RFE3 | Q9RFE3 escherichia |
| 9 | 560 | 24.6 | 413 | Q9RFE3 | Q9RFE3 escherichia |
| 10 | 547 | 24.0 | 413 | Q9RFE3 | Q9RFE3 escherichia |
| 11 | 536 | 23.5 | 414 | Q9RFE3 | Q9RFE3 escherichia |
| 12 | 227.5 | 13.3 | 419 | Q9RFE3 | Q9RFE3 escherichia |
| 13 | 231 | 8.8 | 125 | Q9RFE3 | Q9RFE3 escherichia |
| 14 | 146 | 6.4 | 531 | Q9RFE3 | Q9RFE3 escherichia |
| 15 | 140.5 | 6.2 | 333 | Q9RFE3 | Q9RFE3 escherichia |
| 16 | 137 | 6.0 | 416 | Q9RFE3 | Q9RFE3 escherichia |

| | | | | | | |
|----|-------|-----|-----|---|--------|--------------------|
| 17 | 135 | 5.9 | 423 | 4 | Q9BTU7 | Q9BTU7 homo sapien |
| 18 | 131 | 5.7 | 542 | 5 | Q9U1A2 | Q9U1A2 leishmania |
| 19 | 127 | 5.6 | 395 | 5 | Q9V6B8 | Q9V6B8 drosophila |
| 20 | 126 | 5.5 | 432 | 3 | Q9W2J5 | Q9W2J5 aspergillus |
| 21 | 126 | 5.5 | 465 | 3 | Q9W2J5 | Q9W2J5 aspergillus |
| 22 | 126 | 5.5 | 465 | 3 | Q9W2J5 | Q9W2J5 aspergillus |
| 23 | 126 | 5.5 | 465 | 3 | Q9W2J5 | Q9W2J5 aspergillus |
| 24 | 124.5 | 5.5 | 447 | 5 | Q9U503 | Q9U503 leishmania |
| 25 | 123.5 | 5.4 | 392 | 5 | Q9U503 | Q9U503 drosophila |
| 26 | 123.5 | 5.4 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 27 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 28 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 29 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 30 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 31 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 32 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 33 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 34 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 35 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 36 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 37 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 38 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 39 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 40 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 41 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 42 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 43 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 44 | 120.5 | 5.3 | 447 | 5 | Q9U503 | Q9U503 drosophila |
| 45 | 119.5 | 5.2 | 447 | 5 | Q9U503 | Q9U503 drosophila |

ALIGNMENTS

RESULT 1

| ID | Q9RKE0 | PRELIMINARY: | PRT: | 432 AA. |
|----|---|--------------|------|---------|
| AC | Q9RKE0: | | | |
| DT | 01-JUN-2002 (TRMBLrel. 21, Created) | | | |
| DT | 01-JUN-2002 (TRMBLrel. 21, Last sequence update) | | | |
| DT | 01-JUN-2002 (TRMBLrel. 21, Last annotation update) | | | |
| DE | Periplasmic phosphatidylhydrolase phosphatidylhydrolase. | | | |
| GN | APPA. | | | |
| OS | Escherichia coli. | | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; | | | |
| OC | Escherichia. | | | |
| OX | NCBI_TaxID=562; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=90368616; PubMed=2168385; | | | |
| RA | Dassa J., Marek C., Boquet P.-L.L.; | | | |
| RT | "The complete nucleotide sequence of the Escherichia coli gene appA | | | |
| RT | reveals significant homology between pH 2.5 acid phosphatase and | | | |
| RT | glucosyl-phosphatase." | | | |
| RL | J. Bacteriol. 172:5497-5500(1990). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=94054596; PubMed 1429631; | | | |
| RA | Ostani K., Harms E.H., Stevis P.E., Kuciel K., Zhou M.M., | | | |
| RT | Van Etten R.L.; | | | |
| RT | "Overexpression, site-directed mutagenesis, and mechanism of | | | |
| RT | Escherichia coli acid phosphatase." | | | |
| RL | J. Biol. Chem. 267:22830-22836(1992). | | | |
| DR | EMBL: L03370; AAA00002.1; | | | |
| KW | Hydrolase | | | |
| SQ | SEQUENCE 432 AA, 4697 MW, 583053.5E4692EABF CRC64; | | | |

Query Match 95.5%, Score 2176; BB 2; Length 432;
Post-local Similarity 97.9%; Pred. No. 3.4e-170;
Matches 423; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKALLPFLSLIPLTPQSA.....GFTQIVNEARIPACSLRSH 436
|||||

Db 1 MKAIIIFLSLLIPITQSAAQSEPEPKLESVVIVSABVVRAPTKATQIMODVTPWAMP 60
QY 61 IWPVKAGELTPRGGELIAYLGHYWRQLVADQLLKKQKQPSQVALLAWDERTKIGE 120
Db 61 IWPVKAGELTPRGGELIAYLGHYWRQLVADQLLKKQKQPSQVALLAWDERTKIGE 120
QY 121 AFAAGLAHVAIIVHTQADISSPDLNPDKTGWCOLINANVTUATLIRAGGSADPTGH 180
Db 121 AFAAGLAHVAIIVHTQADISSPDLNPDKTGWCOLINANVTUATLIRAGGSADPTGH 180
QY 181 YQIAFRELERVNFPQSNLTKERKQDSQSLTQALPSSTKVSALVSLTAVSLASMLT 240
Db 181 YQIAFRELERVNFPQSNLTKERKQDSQSLTQALPSSTKVSALVSLTAVSLASMLT 240
QY 241 EFTLLQAGGMPGPGWGRITTSRHWNTLSLINAQFVLLQPTPVARSRAIPLDLTKTA 300
Db 241 EFTLLQAGGMPGPGWGRITTSRHWNTLSLINAQFVLLQPTPVARSRAIPLDLTKTA 300
QY 301 LIPHPQKQAVGVLTPTSVLTACHDINLANIGALFNWTLPGQDPTPGSGELVPERW 360
Db 301 LIPHPQKQAVGVLTPTSVLTACHDINLANIGALFNWTLPGQDPTPGSGELVPERW 360
QY 361 KRLSUNSWTQVSVPTQOMRKPTLSNTPSPKVLITAGFERNAGQMSIAGPTQ 420
Db 361 KRLSUNSWTQVSVPTQOMRKPTLSNTPSPKVLITAGFERNAGQMSIAGPTQ 420
QY 421 LVNEARIPASL 432
Db 421 LVNEARIPASL 432

RESULT 2
QRRK08
ID QRRK08 PRELIMINARY: PROT: 432 AA.
AC QRRK08
DT 01-JUN-2002 (TEMBREl. 21, Created)
DT 01-JUN-2002 (TEMBREl. 21, last sequence update)
DE 01-JUN-2002 (TEMBREl. 21, last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN APPA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID 562;
RN [1]
RX MEDLINE 90368616, PubMed 2164885;
RA Dassa J., Marek C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
reveals significant homology between pH 2.5 acid phosphatase and
glucose-1-phosphatase";
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RX MEDLINE 94054596; PubMed 1429631;
RA Oostanin K., Harms E.H., Steyts P.E., Kuciel P., Zhou M.M.,
Van Elteren R.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
Escherichia coli acid phosphatase";
PL J. Biol. Chem. 267:22840-22846(1992).
DR EMBL: L03373 AAA00004.1;
KW Hydrolase.
SQ SEQUENCE: 432 AA; 46971 MW; 9485530657FCGR45 CR404;

Query Match 95.5%; Score 2176; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 340 170;
Matches 424; conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 MKAIIIFLSLLIPITQSAAQSEPEPKLESVVIVSABVVRAPTKATQIMODVTPWAMP 60
Db 1 MKAIIIFLSLLIPITQSAAQSEPEPKLESVVIVSABVVRAPTKATQIMODVTPWAMP 60
QY 61 IWPVKAGELTPRGGELIAYLGHYWRQLVADQLLKKQKQPSQVALLAWDERTKIGE 120

Db 61 IWPVKAGELTPRGGELIAYLGHYWRQLVADQLLKKQKQPSQVALLAWDERTKIGE 120
QY 121 AFAAGLAHVAIIVHTQADISSPDLNPDKTGWCOLINANVTUATLIRAGGSADPTGH 180
Db 121 AFAAGLAHVAIIVHTQADISSPDLNPDKTGWCOLINANVTUATLIRAGGSADPTGH 180
QY 181 YQIAFRELERVNFPQSNLTKERKQDSQSLTQALPSSTKVSALVSLTAVSLASMLT 240
Db 181 YQIAFRELERVNFPQSNLTKERKQDSQSLTQALPSSTKVSALVSLTAVSLASMLT 240
QY 241 EFTLLQAGGMPGPGWGRITTSRHWNTLSLINAQFVLLQPTPVARSRAIPLDLTKTA 300
Db 241 EFTLLQAGGMPGPGWGRITTSRHWNTLSLINAQFVLLQPTPVARSRAIPLDLTKTA 300
QY 301 LIPHPQKQAVGVLTPTSVLTACHDINLANIGALFNWTLPGQDPTPGSGELVPERW 360
Db 301 LIPHPQKQAVGVLTPTSVLTACHDINLANIGALFNWTLPGQDPTPGSGELVPERW 360
QY 361 KRLSUNSWTQVSVPTQOMRKPTLSNTPSPKVLITAGFERNAGQMSIAGPTQ 420
Db 361 KRLSUNSWTQVSVPTQOMRKPTLSNTPSPKVLITAGFERNAGQMSIAGPTQ 420
QY 421 LVNEARIPASL 432
Db 421 LVNEARIPASL 432

RESULT 3
QRRK07
ID QRRK07 PRELIMINARY: PROT: 432 AA.
AC QRRK07
DT 01-JUN-2002 (TEMBREl. 21, Created)
DT 01-JUN-2002 (TEMBREl. 21, last sequence update)
DE 01-JUN-2002 (TEMBREl. 21, last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN APPA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID 562;
RN [1]
RX MEDLINE 90368616; PubMed 2164885;
RA Dassa J., Marek C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
reveals significant homology between pH 2.5 acid phosphatase and
glucose-1-phosphatase";
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RX MEDLINE 94054596; PubMed 1429631;
RA Oostanin K., Harms E.H., Steyts P.E., Kuciel P., Zhou M.M.,
Van Elteren R.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
Escherichia coli acid phosphatase";
PL J. Biol. Chem. 267:22840-22846(1992).
DR EMBL: L03373 AAA00005.1;
KW Hydrolase.
SQ SEQUENCE: 432 AA; 46971 MW; 750954304AGP14AG CR464;

Query Match 95.5%; Score 2176; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 340 170;
Matches 424; conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 MKAIIIFLSLLIPITQSAAQSEPEPKLESVVIVSABVVRAPTKATQIMODVTPWAMP 60
Db 1 MKAIIIFLSLLIPITQSAAQSEPEPKLESVVIVSABVVRAPTKATQIMODVTPWAMP 60
QY 61 IWPVKAGELTPRGGELIAYLGHYWRQLVADQLLKKQKQPSQVALLAWDERTKIGE 120
Db 61 IWPVKAGELTPRGGELIAYLGHYWRQLVADQLLKKQKQPSQVALLAWDERTKIGE 120

QY 121 AFAAGLAPDCAITVHTQADTSPPDLPNPI KTCVGLDNDANVTDAILEPAGSSTADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSPPDLPNPI KTCVGLDNDANVTDAILEPAGSSTADFTGH 180
QY 181 YQATFRELKRVNFPQSNLCKLKKQDESSLSLQALFSELKVSADVYVSTGAVVASHMLT 240
DB 181 YQATFRELKRVNFPQSNLCKLKKQDESSLSLQALFSELKVSADVYVSTGAVVASHMLT 240
QY 241 EITFLQQAQGMPEFGWGRITLSDHWNLLSLHNAACVLLAGFTFVAFSEKATLFLTKTA 300
DB 241 EITFLQQAQGMPEFGWGRITLSDHWNLLSLHNAACVLLAGFTFVAFSEKATLFLTKTA 300
QY 301 LIPHPKQKAYGVLTFTSVLF TAGHDTNLNLGCALELNNWTLPGQDNTDPPGGELVFFPW 360
DB 301 LIPHPKQKAYGVLTFTSVLF TAGHDTNLNLGCALELNNWTLPGQDNTDPPGGELVFFPW 360
QY 361 RLSDNSQWIVQSVLFTQGMKDKTPLSLNTPPGVRKTLTACCERNAGGMSLAGFTQ 420
DB 361 RLSDNSQWIVQSVLFTQGMKDKTPLSLNTPPGVRKTLTACCERNAGGMSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 4
Q8RKD9
ID Q8RKD9 PRELIMINARY: PRT: 432 AA.
AC Q8RKD9
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN APPA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID:562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9368616; PubMed=2168385;
RA Basso J., Marck C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94054596; PubMed 1429631;
RA Ostlund K., Hains E.H., Stevis E.E., Kacieli E., Zhou M.M.,
Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase".
RL J. Biol. Chem. 267:22830-22836(1992).
DR EMBL: L03371; AAA000003.1; -;
KW Hydrolase.
SQ SEQUENCE 432 AA; 46901 MW; AF96041FAA613AA05 CRef64;

Query Match 95.5%; Score 2176; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 340-170;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKAALLPELSLLILPISQSAFAQSEPELKLSSWIVSPHGVAPATKATQLMQVTPDAMP 60
DB 1 MKAALLPELSLLILPISQSAFAQSEPELKLSSWIVSPHGVAPATKATQLMQVTPDAMP 60
QY 61 TWPKLGLWLTDSOGELIATVHTQADTSPPDLPNPI KTCVGLDNDANVTDAILEPAGSSTADFTGH 120
DB 61 TWPKLGLWLTDSOGELIATVHTQADTSPPDLPNPI KTCVGLDNDANVTDAILEPAGSSTADFTGH 120
QY 121 AFAAGLAPDCAITVHTQADTSPPDLPNPI KTCVGLDNDANVTDAILEPAGSSTADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSPPDLPNPI KTCVGLDNDANVTDAILEPAGSSTADFTGH 180

QY 181 YQATFRELKRVNFPQSNLCKLKKQDESSLSLQALFSELKVSADVYVSTGAVVASHMLT 240
DB 181 YQATFRELKRVNFPQSNLCKLKKQDESSLSLQALFSELKVSADVYVSTGAVVASHMLT 240
QY 241 EITFLQQAQGMPEFGWGRITLSDHWNLLSLHNAACVLLAGFTFVAFSEKATLFLTKTA 300
DB 241 EITFLQQAQGMPEFGWGRITLSDHWNLLSLHNAACVLLAGFTFVAFSEKATLFLTKTA 300
QY 301 LIPHPKQKAYGVLTFTSVLF TAGHDTNLNLGCALELNNWTLPGQDNTDPPGGELVFFPW 360
DB 301 LIPHPKQKAYGVLTFTSVLF TAGHDTNLNLGCALELNNWTLPGQDNTDPPGGELVFFPW 360
QY 361 RLSDNSQWIVQSVLFTQGMKDKTPLSLNTPPGVRKTLTACCERNAGGMSLAGFTQ 420
DB 361 RLSDNSQWIVQSVLFTQGMKDKTPLSLNTPPGVRKTLTACCERNAGGMSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 5
Q8RKD9
ID Q8RKD9 PRELIMINARY: PRT: 432 AA.
AC Q8RKD9
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN APPA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID:562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9368616; PubMed=2168385;
RA Basso J., Marck C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94054596; PubMed=1429631;
RA Ostlund K., Hains E.H., Stevis E.E., Kacieli E., Zhou M.M.,
Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase".
RL J. Biol. Chem. 267:22830-22836(1992).
DR EMBL: L03371; AAA000003.1; -;
KW Hydrolase.
SQ SEQUENCE 432 AA; 47033 MW; 0124B93FAC2E8175 CRef64;

Query Match 95.4%; Score 2175; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 410-170;
Matches 423; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 1 MKAALLPELSLLILPISQSAFAQSEPELKLSSWIVSPHGVAPATKATQLMQVTPDAMP 60
DB 1 MKAALLPELSLLILPISQSAFAQSEPELKLSSWIVSPHGVAPATKATQLMQVTPDAMP 60
QY 61 TWPKLGLWLTDSOGELIATVHTQADTSPPDLPNPI KTCVGLDNDANVTDAILEPAGSSTADFTGH 120
DB 61 TWPKLGLWLTDSOGELIATVHTQADTSPPDLPNPI KTCVGLDNDANVTDAILEPAGSSTADFTGH 120
QY 121 AFAAGLAPDCAITVHTQADTSPPDLPNPI KTCVGLDNDANVTDAILEPAGSSTADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSPPDLPNPI KTCVGLDNDANVTDAILEPAGSSTADFTGH 180
QY 181 YQATFRELKRVNFPQSNLCKLKKQDESSLSLQALFSELKVSADVYVSTGAVVASHMLT 240
DB 181 YQATFRELKRVNFPQSNLCKLKKQDESSLSLQALFSELKVSADVYVSTGAVVASHMLT 240

```

Db 181 ROTAFRELERVINFPGSNLCITKRKQKQSFCSLTQALPSEKLVKSDNVSLTQAVSLASMIT 240
QY 241 EFTLLQAQAQMPFPGWGRITISHQWNTLTLISLNAQFILLQRTPEVARSRAIPLLDLIKTA 300
Db 241 EFTLLQAQAQMPFPGWGRITISHQWNTLTLISLNAQFILLQRTPEVARSRAIPLLDLIKTA 300
QY 301 LTPHPQKQAYGVLTPTSVLFIAGHDTNLANIAGALELNWLTLPQDNTPPGGEIVFERW 360
Db 301 LTPHPQKQAYGVLTPTSVLFIAGHDTNLANIAGALELNWLTLPQDNTPPGGEIVFERW 360
QY 361 RRLSNSQWLTQVSLVFTLQOMRKTPLSLNTPPGEVKLTLAGTEERNACQMCSLAGTQ 420
Db 361 RRLSNSQWLTQVSLVFTLQOMRKTPLSLNTPPGEVKLTLAGTEERNACQMCSLAGTQ 420
QY 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432
RESULT 6
Q8RKJ5
ID Q8RKJ5 PRELIMINARY: PRT: 432 AA.
AC Q8RKJ5
DT 01-JUN-2002 (Tremblrel, 21, Created)
DT 01-JUN-2002 (Tremblrel, 21, Last sequence update)
DE periplasmic phosphoanhydride phosphatase.
GN APFA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
PN 111
RP SEQUENCE FROM N.A.
RX MEDLINE 90468616; PubMed=2168185;
RA Dassa J., Marek C., Boquet P., Leli.
RT "The complete nucleotide sequence of the Escherichia coli gene apfa
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RI J. Bacteriol. 172:5497-5500 (1990).
RN 121
RP SEQUENCE FROM N.A.
RX MEDLINE 94054594; PubMed=1429643;
RA Ostlund K., Bains R.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase."
RI J. Biol. Chem. 267:22840-22846 (1992).
DR EMBL: L03375; AAA00007.1;
KW Hydrolyase.
SQ SEQUENCE 432 AA: 46990 MW: 951F304EA9A1A47C 6RC64;
Query Match 95.49, Score 2172, ID 2, Length 432.
Best Local Similarity 97.98, Pred. No. 7,26-170;
Matches 42%, Conservative 0, Mismatches 9, Indels 0, Gaps 0;
QY 1 MKAILTFELSLILITPUSAFQAQSEPELKI ESNVIVSHPHVPATKATLQMGVTPDAMP 60
Db 1 MKAILTFELSLILITPUSAFQAQSEPELKI ESNVIVSHPHVPATKATLQMGVTPDAMP 60
QY 61 TWPVKRLDELTRGCHT TAYICHYWRPPLVAGSLPKKQSTSSGQAVIALVIEPIKTEGE 120
Db 61 TWPVKRLDELTRGCHT TAYICHYWRPPLVAGSLPKKQSTSSGQAVIALVIEPIKTEGE 120
QY 121 AFAACIAPVAVI TIVTQADTSSDQIPNPDKTQVQIDNANVTDAIIPPAAGSTADPTCH 180
Db 121 AFAACIAPVAVI TIVTQADTSSDQIPNPDKTQVQIDNANVTDAIIPPAAGSTADPTCH 180
QY 181 YOTAFRELERVINFPGSNLCITKRKQKQSFCSLTQALPSEKLVKSDNVSLTQAVSLASMIT 240
Db 181 YOTAFRELERVINFPGSNLCITKRKQKQSFCSLTQALPSEKLVKSDNVSLTQAVSLASMIT 240
QY 241 EFTLLQAQAQMPFPGWGRITISHQWNTLTLISLNAQFILLQRTPEVARSRAIPLLDLIKTA 300

```

```

Db 241 EFTLLQAQAQMPFPGWGRITISHQWNTLTLISLNAQFILLQRTPEVARSRAIPLLDLIKTA 300
QY 301 LTPHPQKQAYGVLTPTSVLFIAGHDTNLANIAGALELNWLTLPQDNTPPGGEIVFERW 360
Db 301 LTPHPQKQAYGVLTPTSVLFIAGHDTNLANIAGALELNWLTLPQDNTPPGGEIVFERW 360
QY 361 RRLSNSQWLTQVSLVFTLQOMRKTPLSLNTPPGEVKLTLAGTEERNACQMCSLAGTQ 420
Db 361 RRLSNSQWLTQVSLVFTLQOMRKTPLSLNTPPGEVKLTLAGTEERNACQMCSLAGTQ 420
QY 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432
RESULT 7
Q8XC29
ID Q8XC29 PRELIMINARY: PRT: 444 AA.
AC Q8XC29
DT 01-MAR-2002 (Tremblrel, 20, Created)
DT 01-MAR-2002 (Tremblrel, 20, Last sequence update)
DE Phosphoanhydride phosphatase, pH 2.5 acid phosphatase,
DE periplasmic.
GN APFA OR 21397 OR EGS1136.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
PN 111
RP SEQUENCE FROM N.A.
RX STRAIN O157:H7 / EHL933 / ATCC 700927;
RA Medina N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Green J., Kirkpatrick H.A.,
RA Kosfai G., Hackett J., Klunk S., Boutin A., Shao Y., Miller L.,
RA Grobeck J.J., Davis N.W., Lim A., Bimblant E.J., Potamous K.,
RA Apodaca J., Anantharaman L.S., Lin J., Yen G., Schwartz D.C.,
RA Welch K.A., Jattoor F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RI Nature 409:529-533 (2001).
RN 121
RP SEQUENCE FROM N.A.
RX STRAIN O157:H7 / KMO 0609952;
RX MEDLINE 21156241; PubMed 11250799;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Take T.,
RA Tada T., Takami H., Honda T., Sasakiwa C., Ogasawara N., Yasumura T.,
RA Kubara S., Shiba T., Hattori M., Shimaqawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RI DNA Res. 8:31-22 (2001).
DR EMBL: AF005292; AA055528.1; AL1_1111.
DR EMBL: AP002554; HA04559.1;
DR InterPro: IPR000560; HIsAc_phosphatase.
DR Pfam: PF00426; acid_phosphat_1.
DR PROSITE: PS00616; HIsAcID_PHSPHAT_1; 1.
DR PROSITE: PS00778; HIsAcID_PHSPHAT_2; 1.
KW Complete proteome.
SQ SEQUENCE 444 AA: 47437 MW: F192DF2B1E64F964 29C64;
Query Match 94.19, Score 2144, ID 19, Length 444.
Best Local Similarity 96.58, Pred. No. 1,40-167;
Matches 41%, Conservative 0, Mismatches 14, Indels 2, Gaps 1;
QY 1 MKAILTFELSLILITPUSAFQAQSEPELKI ESNVIVSHPHVPATKATLQMGVTPDAMP 60
Db 1 MKAILTFELSLILITPUSAFQAQSEPELKI ESNVIVSHPHVPATKATLQMGVTPDAMP 60
QY 61 TWPVKRLDELTRGCHT TAYICHYWRPPLVAGSLPKKQSTSSGQAVIALVIEPIKTEGE 120
Db 61 TWPVKRLDELTRGCHT TAYICHYWRPPLVAGSLPKKQSTSSGQAVIALVIEPIKTEGE 120

```


QY 13 LHTQSNAPSPF-LHTFVWIVSGVGPAP-TPATQLMKGVTDAMTPWVKLGEL 70
 DB 13 VVLLASNAAGIVPQVQGLQVAVIMWSGNGEVAFAANAEVSGSTPRKWFLEWVWGGCL 72
 QY 71 PRGGLIAYLGHYGRQRLVAGGLPKQVTPQSGAVATADVETPKTSFAFAA-GIAPDC 130
 DB 71 LHTQSNAPSPF-LHTFVWIVSGVGPAP-TPATQLMKGVTDAMTPWVKLGEL 70
 QY 73 TKGAVI FVYMGMYPWGLAFGLMVSSEFPPFGVYAYANSLSLQFVALAAGFFIIGAFPGG 132
 DB 73 TKGAVI FVYMGMYPWGLAFGLMVSSEFPPFGVYAYANSLSLQFVALAAGFFIIGAFPGG 132
 QY 141 ATTHTQADTSPPDLPNPKTKVQVQINAVTDG---AIIIPRAGSIAIDTGHVGTAPPP 187
 DB 141 ATTHTQADTSPPDLPNPKTKVQVQINAVTDG---AIIIPRAGSIAIDTGHVGTAPPP 187
 QY 143 DIPVHQKMKMTDPTNFVITD-----GSAAFSEGAVAAMEKESKI-QUTHSYGL----- 183
 DB 143 DIPVHQKMKMTDPTNFVITD-----GSAAFSEGAVAAMEKESKI-QUTHSYGL----- 183
 QY 188 LERVINEQSNICPKPKQFQFSSILTALPSELKVSALPVSILGAVSLASMIITFIFLQD 247
 DB 188 LERVINEQSNICPKPKQFQFSSILTALPSELKVSALPVSILGAVSLASMIITFIFLQD 247
 QY 184 LERVINEQSNICPKPKQFQFSSILTALPSELKVSALPVSILGAVSLASMIITFIFLQD 247
 DB 184 LERVINEQSNICPKPKQFQFSSILTALPSELKVSALPVSILGAVSLASMIITFIFLQD 247
 QY 248 AQSMF---PPWPTDSQWNTTTSINAGFELGQRTFPAFSAFALPGLALIKIALIDEP 305
 DB 248 AQSMF---PPWPTDSQWNTTTSINAGFELGQRTFPAFSAFALPGLALIKIALIDEP 305
 QY 240 VYSGEMFVAVHFKPSFQWVLSKIKNGYFUSFTSEVAVNVAKPLVSYLPAAL 295
 DB 240 VYSGEMFVAVHFKPSFQWVLSKIKNGYFUSFTSEVAVNVAKPLVSYLPAAL 295
 QY 206 PQKAVAVVITPS---VIFACHUNIANQNALPLN-WTIDGQNTPTFQSEIVFPPW 361
 DB 206 PQKAVAVVITPS---VIFACHUNIANQNALPLN-WTIDGQNTPTFQSEIVFPPW 361
 QY 206 -----VIDTSAKFTIVVHNSNTASTITATIDKRYQIHQGNERTFEGSKIVFPPW 345
 DB 206 -----VIDTSAKFTIVVHNSNTASTITATIDKRYQIHQGNERTFEGSKIVFPPW 345
 QY 362 PLSNSQWVAVSVFATQSCMPDPTDSTNTPPFVZITLAVGEPFPAKQWMSAGPTG 401
 DB 362 PLSNSQWVAVSVFATQSCMPDPTDSTNTPPFVZITLAVGEPFPAKQWMSAGPTG 401
 QY 349 DSPKAPQIMKIPVYVYQSNQFPAQVATLQAPAGVETLFSFGVQFVAGVQFVLSV 407
 DB 349 DSPKAPQIMKIPVYVYQSNQFPAQVATLQAPAGVETLFSFGVQFVAGVQFVLSV 407
 QY 422 VNEA 425
 DB 408 LNEA 411

RESULT 10
 Q9AAQ4 PRELIMINARY; PRI: 413 AA.
 DT 01-MAR-2002 (TREMblrel, 20, Created)
 DT 01-MAR-2002 (TREMblrel, 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)
 DE Glucose-1 phosphatase (G1Pase); secreted.
 GN STY1153.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Salmonella.
 CC NCB1_TaxID 601;
 RN (1)
 SEQUENCE FROM N.A.
 RC STRAIN CT18;
 RX MEDLINE: 21534947; PubMed: 11677608;
 RA Parkhill J., Isouard G., James K.D., Thomson N.R., Pickard D., Watin J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Cornet P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltham J., Hamlin N., Haque A., Hien T., Holtroyd S., Jagels K.,
 RA Karp A., Karsen J.S., Leather S., Meele S., O'Garra P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skellern J., Stevens K.,
 RA Whitehead S., Worroll H.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18";
 RL Nature 413:848-852(2001).
 RL EMBL: AL527269; CA008242 1;
 DB InterPro: IPR000560; HisAc_phosphatase.
 DB Pfam: PF00328; acid_phosphat_1;
 DB PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1;
 DB PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1;
 KW Complete proteome.
 SQ SEQUENCE: 413 AA; 45630 MW; E9FE563CHRRD9D26 CR64;
 Query Match 24.0%; Score 547; DB 16; Length 413;
 Best Local Similarity 44.7%; Pred. No. 1,4+46;
 Matches 148; Conservative 66; Mismatches 181; Indels 24; Gaps 10;

RESULT 11

Q9AAQ4

PRELIMINARY;

PRI: 414 AA.

DT 01-JUN-2001 (TREMblrel, 17, Created)

DT 01-JUN-2001 (TREMblrel, 17, Last sequence update)

DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)

DE periplasmic phosphoanhydride phosphohydrolase.

GN CC0542.

OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;

CC Caulobacter.

CC NCB1_TaxID 155892;

RN (1)

SEQUENCE FROM N.A.

RC STRAIN-A10C 9089 / CR15;

RX MEDLINE: 21174698; PubMed: 11259647;

RA Nierman W.C., Feldblyum T.V., Lamb M.T., Paulsen O., Nelson K.E.,

RA Eisen J., Heidelberg J.F., Alley M.R.K., Gaitanaris G., McKee K.L.,

RA Potocka L., Nelson W.C., Newton A., Stephens C., Hanks N.D., Ely R.,

RA Deboy R.T., Dodson R.J., Burkin A.S., Gwin M.L., Haft J.H.,

RA Kileyay J.F., Smit J., Graven M.B., Khouri B., Shetty B., Portz G.,

RA Otterback T., Tran K., Wolf A., Vamathevan J., Ermolova M., White O.,

RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;

RT "Complete genome sequence of Caulobacter crescentus";

RL Proc. Natl. Acad. Sci. U.S.A. 98:4140-4141(2001).

DB EMBL: A9009227; AAF2529.1; 1;

DB BSRP: ps7102; 10K1.

DB TIGR: CR0542;

DR InterPro: IPR00560; HisAc_phosphatase.

DR Pfam: PF00328; acid_phosphat_1;

DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1;

DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1;

KW Hydrolyase; Complete proteome.

SQ SEQUENCE: 414 AA; 44279 MW; 7105b64EA01c4e81 CR64;

Query Match 24.5%; Score 546; DB 16; Length 414;

Best Local Similarity 44.6%; Pred. No. 1,4+46;

Matches 145; Conservative 62; Mismatches 166; Indels 49; Gaps 11;

QY 9 LSL---LTPITPSAFAQSEFPELEKESVIVSPFPVPAITATQALMQAVIHTAWIHWVR 65

DB 9 LSL---LTPITPSAFAQSEFPELEKESVIVSPFPVPAITATQALMQAVIHTAWIHWVR 65

DB 3 LSLRTAVVILALASAGASAE-TLEKVVILSRHVPKSMSSPELEFAGAGIHWRFVVP 60

| | | | |
|-----------|-------------|---|---------|
| QY | 247 | QACMP--EPWCPITTSWNTTISLHNAQFLLOPTEVAASVATPLHDLIKTALPH | 404 |
| DB | 232 | WSPNPLSLQAWCKITQAPQITALLPIITENYNINMIVVTAAGAGSVITNAMIQVRIE | 291 |
| QY | 305 | PPQKAYGVTLPTSVLFIAGHPTNLNLGSALEFINNTIET--QPDNTPAGELVFERWR | 362 |
| DB | 292 | ANDNVW-----LLLVADNTIAMVPTIMNFNSWQIDYSGNIPPGSSVLLEWRTR | 444 |
| QY | 363 | LSNSQWLVSVLVEQI--LQMPPTETSLNTPGSEVEL--LAEFERNAQJMC-- | 414 |
| DB | 344 | KAENA--LCGSIFQAQGLDDI-----ELQITPAQHPMLEPTEWRGPTETQIVNLTQEP | 496 |
| QY | 414 | --SLACFTQIVNEARIPACSL | 432 |
| DB | 396 | QAATTAIGQIDRSSAPAVAM | 416 |
| RESULT 13 | | | |
| Q9BZG2 | | Q9BZG2 | 426 AA. |
| AC | Q9BZG2 | Q9BZG2 | 426 AA. |
| AC | Q9BZG2 | Q9BZG2 | 426 AA. |
| DT | 01-JUN-2001 | (TREMBLrel 17, Created) | |
| DT | 01-JUN-2001 | (TREMBLrel 17, Last sequence update) | |
| DT | 01-MAP-2002 | (TREMBLrel 20, Last annotation update) | |
| GN | | Acid phosphatase. | |
| DE | | ACPT. | |
| OS | | Homo sapiens (Human) | |
| OC | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| OC | | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| OX | | NCBI_TaxID=9606; | |
| RN | | [1] | |
| RP | | SEQUENCE FROM N.A. | |
| RP | | MEDLINE=21309073; PubMed=11414767; | |
| PT | | Yusef Q M, Diamandis M, Jung K, Diamandis E.F.; | |
| PT | | "Molecular cloning of a novel human acid phosphatase gene (ACPT) that | |
| RT | | is highly expressed in the testis."; | |
| FL | | Genomics 74:385-395(2001) | |
| DB | | EMBL; AF321918; AAK0393.1; -- | |
| DB | | HSST; F15303; 2HPA. | |
| DB | | InterPro; IPR000560; HisAc-phsptase | |
| DB | | Pfam; PF00328; acid.phosphat; 1. | |
| DB | | PROSITE; PS00616; HIS.ACID.PHOSPHAT.1; | |
| DB | | PROSITE; PS00778; HIS.ACID.PHOSPHAT.2; UNKNOWN.1. | |
| SO | | SEQUENCE 426 AA; 4509 MW; RS304580418061_CPC64; | |

```

Query Match      8.8%; Score 201, DB 4; Length 426;
Best Local Similarity 26.5%; Pred.No.3.4e-08;
Matches 119; Conservative 53; Mismatches 175; Indels 102; Gaps
QY      7 PFLSLPLTLPQSAFAQSPFLKLESWIVSRGVRAPTKATQMLQGVTPDPAWTFIVKL 66
QL      13 PFLSLPLTLPQSAFAQSPFLKLESWIVSRGVRAPTKATQMLQGVTPDPAWTFIVKL 67
QY      67 GFLTPDGGTITAVYGHYWPQPIVADGCIIRGQSPQSQDAVALIAVDEPTPTVFAVAGLL 126
QL      126 GFLTPDGGTITAVYGHYWPQPIVADGCIIRGQSPQSQDAVALIAVDEPTPTVFAVAGLL 127
QY      68 GFLTPDGGTITAVYGHYWPQPIVADGCIIRGQSPQSQDAVALIAVDEPTPTVFAVAGLL 127
QL      127 GFLTPDGGTITAVYGHYWPQPIVADGCIIRGQSPQSQDAVALIAVDEPTPTVFAVAGLL 128
QY      127 APLCALIVHLCATSSPPPLFNELKTCVQVQDPAWTVDAILEPAGSSIAFTGHYQIAPR 180
QL      180 APLCALIVHLCATSSPPPLFNELKTCVQVQDPAWTVDAILEPAGSSIAFTGHYQIAPR 181
QY      124 PTEAA-----PGSPFPAPPIPVH-----TVPVAE----- 148
QL      148 PTEAA-----PGSPFPAPPIPVH-----TVPVAE----- 149
QY      187 FTFRVNFQSNIC-----IKRRQDESCSLTALFLKVSACVSLIGAVSL----- 245
QL      245 FTFRVNFQSNIC-----IKRRQDESCSLTALFLKVSACVSLIGAVSL----- 246
QY      149 DLLKLFPMRS-CPRYHELIRE APLAAHYQLALECWTGCTSLKLENFTG-LSLVGEPLR 201
QL      201 DLLKLFPMRS-CPRYHELIRE APLAAHYQLALECWTGCTSLKLENFTG-LSLVGEPLR 202
QY      235 -ASMLJEIPLLLQACMPFPQWCTITDSSHAWNTELLSHUNAGFELLQRTPEVARSATPLL 294
QL      294 -ASMLJEIPLLLQACMPFPQWCTITDSSHAWNTELLSHUNAGFELLQRTPEVARSATPLL 295
QY      204 FAKKVIPTLMCCAGCLPTIPAWA-----SPDWLPTL--AQLEALNLCAGVCPPPAAEKA 255
QL      255 FAKKVIPTLMCCAGCLPTIPAWA-----SPDWLPTL--AQLEALNLCAGVCPPPAAEKA 256
QY      295 EI-----IKTALTPHPDQKQAGVGLTPTSVLFIAGHTYNTANIIGALEHNLPLQNPINTP 450
QL      450 EI-----IKTALTPHPDQKQAGVGLTPTSVLFIAGHTYNTANIIGALEHNLPLQNPINTP 451
QY      256 QITGCTGIIINAIANFSAPVDIG--LPLEKMYSAHNTSLALACVLSLGL-----VYCHTP 407
QL      407 QITGCTGIIINAIANFSAPVDIG--LPLEKMYSAHNTSLALACVLSLGL-----VYCHTP 408

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 15:52:40 : Search time 21 seconds
(without alignments)
1995.934 Million cell updates/sec

Title: US-09-866-379A-10
Perfect score: 2279
Sequence: 1 MEALLFFSLTLLTQSA..... : RPTGVVREAHDA:ELFSLH: 43C

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR23*

1: pir1*

2: pir2*

3: pir3*

4: pir4*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 2182 | 95.7 | 422 | B36733 | acid phosphatase (|
| 2 | 2144 | 94.1 | 434 | B00776 | phosphonhydride F |
| 3 | 2144 | 94.1 | 444 | B85633 | hypothetical prote |
| 4 | 959.5 | 42.1 | 441 | A20201 | acid phosphatase (|
| 5 | 560 | 24.6 | 413 | B36773 | periplasmic glucos |
| 6 | 560 | 24.6 | 413 | B85635 | periplasmic glucos |
| 7 | 554 | 24.3 | 413 | JN0087 | glucose-1-phosphat |
| 8 | 547 | 24.0 | 413 | A20532 | glucose-1-phosphat |
| 9 | 538 | 23.6 | 417 | S25227 | glucose-1-phosphat |
| 10 | 536 | 23.5 | 414 | B87316 | periplasmic phosph |
| 11 | 137 | 6.0 | 416 | B16358 | hypothetical prote |
| 12 | 134.5 | 5.9 | 423 | S85167 | acid phosphatase (|
| 13 | 119 | 5.3 | 433 | A23395 | acid phosphatase (|
| 14 | 112.5 | 4.9 | 428 | S64682 | acid phosphatase (|
| 15 | 111.5 | 4.9 | 421 | S11712 | acid phosphatase (|
| 16 | 109.5 | 4.8 | 344 | B87130 | protein F52E1.8 (1 |
| 17 | 109.5 | 4.8 | 457 | A56925 | paired box transcr |
| 18 | 109.5 | 4.8 | 479 | JN0890 | acid phosphatase (|
| 19 | 108 | 4.7 | 396 | JH0610 | acid phosphatase (|
| 20 | 107.5 | 4.7 | 459 | S53553 | paired box transcr |
| 21 | 107.5 | 4.7 | 479 | JN0715 | 3-phytase (EC 3.1 |
| 22 | 107 | 4.7 | 526 | S40450 | tyrosine receptor |
| 23 | 105 | 4.6 | 453 | A54429 | paired box transcr |
| 24 | 104.5 | 4.6 | 397 | C81716 | hypothetical prote |
| 25 | 103.5 | 4.5 | 537 | S54770 | secreted acid phos |
| 26 | 103.5 | 4.5 | 888 | S45726 | secreted acid phos |
| 27 | 103.5 | 4.5 | 1048 | EV858C | exonuclease (EC 3. |
| 28 | 103 | 4.5 | 769 | B87e81 | tyrosine kinase li |
| 29 | 101.5 | 4.5 | 356 | I15594 | hypothetical prote |

| | | | | | | |
|----|-------|-----|------|---|--------|----------------------|
| 30 | 101.5 | 4.5 | 635 | 2 | A45266 | MPL-P protein proc |
| 31 | 101.5 | 4.5 | 789 | 2 | AE0064 | DNA-directed DNA p |
| 32 | 101.5 | 4.5 | 1214 | 2 | A52897 | conserved hypofact |
| 33 | 101.5 | 4.5 | 1387 | 2 | A97673 | probable periplasm |
| 34 | 100.5 | 4.4 | 425 | 2 | D83186 | hypothetical prote |
| 35 | 100 | 4.4 | 971 | 2 | T32883 | hypothetical prote |
| 36 | 99.5 | 4.4 | 234 | 2 | F42696 | thrombin (EC 3.4.4.2 |
| 37 | 99 | 4.3 | 560 | 2 | P84341 | hypothetical prote |
| 38 | 99 | 4.3 | 969 | 2 | T27997 | hypothetical prote |
| 39 | 99 | 4.3 | 1413 | 2 | D88844 | protein ZK742.1 (1 |
| 40 | 98.5 | 4.3 | 381 | 2 | JH0152 | acid phosphatase (|
| 41 | 98 | 4.3 | 1034 | 2 | AB0551 | exonuclease Sber 1 |
| 42 | 97.5 | 4.3 | 4056 | 2 | H96599 | protein F14J16.10 |
| 43 | 96.5 | 4.2 | 228 | 2 | A83572 | pyridoxal phosphat |
| 44 | 96.5 | 4.2 | 374 | 2 | I39781 | subtilisin (EC 3.4 |
| 45 | 96.5 | 4.2 | 1047 | 2 | G90684 | ATP-dependent dsDN |

ALIGNMENTS

RESULT 1

B36733

acid phosphatase (EC 3.1.3.2) precursor - Escherichia coli (strain K-12)
N.Altitude bases, phosphonhydride phosphatase, periplasmic, pyruvate 6
C.Species: Escherichia coli
C.Date: 19-Apr-1991 #sequence_revision: 19-Apr-1991 #excl_change: 31-Mar-2002
C.Accession: B36733; S18018; B64839; A26534; S17960; S33278
R.Dassa, J.; Marck, C.; Boquet, P.L.
J. Bacteriol. 172, 5497-5500, 1990
A.Title: The complete nucleotide sequence of the Escherichia coli gene *appA* reveals
A.Reference number: A36733; MUID:90368616; PMID:2168485
A.Accession: B36733

A>Status: preliminary

A.Molecule type: DNA

A.Cross-references: 1-432 <DAS>

A.Cross-references: op M59708; NID:Q145283; PIRN:AA72086.1; PIR:Q145285

P.Greiner, P. J. Jany, K. D

Biol. Chem. Hoppe-Seyler 372, 664-665, 1991

A.Title: Characterization of a phytase from Escherichia coli.

A.Reference number: S18018

A.Accession: S18018

A.Molecule type: protein

A.Residues: 23-33 <SORE>

P.Plattner, F.R.; Maub, R.; Shao, Y.

Science 277, 1453-1462, 1997

A.Title: The complete genome sequence of Escherichia coli K-12.

A.Reference number: A64726; MUID:97426617; PMID:9278503

A.Accession: B64839

A>Status: nucleic acid sequence not shown; translation not shown

A.Molecule type: DNA

A.Residues: 1-432 <BLAT>

A.Cross-references: GB:AE002002; GB:000026; NID:Q1367111; PIRN:AA74065.1; PIR:Q136711

A.Experimental Source: Strain K-12, Substrain MG1655

P.Touati, E.; Danchin, A.

Biochimie 69, 215-221, 1997

A.Title: The structure of the promoter and amino terminal region of the *prf* 2.5 gene

A.Reference number: A26534; MUID:97271766; PMID:9039201

A.Accession: A26534

A.Molecule type: DNA

A.Residues: 1-50; NAGGCHPRMANAG/65; T/67-74; DV/77-111; S/ <TOD>

A.Cross-references: GB:X05471; NID:Q136925; PIRN:CAA24031.1; PIR:Q136927

F.Dassa, J.; Fathi, H.; Marck, C.; Bion, M.; Klotter-Bontemps, M.; Roquet, P.L.

Mol. Gen. Genet. 229, 341-352, 1991

A.Title: A new *oxyS*-regulated operon in Escherichia coli comprises the genes for

A.Accession: S17960

A>Status: not compared with conceptual translation

A.Molecule type: DNA

A.Residues: 1-17 <DAS>

A.Cross-references: GB:SC3811; NID:Q286556; PIRN:AA620286.1; PIR:Q238659

P.Greiner, P.; Klotter-Bontemps, M.; Jany, K.D.

[illegible]

RESULT 8
 AG0642
 glucose-1-phosphatase precursor (c1pase), secreted (imported)
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 06/06/2001 15:40:00
 C:Accession: AG0642
 P:Patrizi, L.; Bonaldi, G.; James, K.D.; Thomson, N.K.; Packard, D.; Wain, L.; Chakr-
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, K.M.; Dwyer, L.; Whittam, N.; Par-
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rothbard, K.; Simmonds, M.; Skellern, J.; Theodor-
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AG0642
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1413 x PAR
 A:Cross-references: GB:AL133482; FIDN:AF0642.1; FIDN:AF0642.2
 C:Qualities:
 A:Gene: STY1153
 Query Match 24.09% Score 547 Length 432
 Best Local Similarity 33.78% Prev. Node 46

Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:01 11 November 2014

1

Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:51 11 November 2014

C:Accession: A33395
R:Himeno, M.; Fujita, H.; Nomachi, Y.; Kono, A.; Kato, K.
Biochem. Biophys. Res. Commun. 162, 1044-1053, 1989
A:Title: Isolation and sequencing of a cDNA clone encoding acid phosphatase in rat liver
A:Reference number: A33395; MUID:86250910; PMID:2764916
A:Accession: A33395
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-423 <SHM>
A:Cross-references: GB:827693, EMBL:344174 1; FID:930934
C:Superfamily: mammalian acid phosphatase
C:Keywords: phosphatidase; phosphoprotein; phosphoric monoester hydrolase
F:41/Active site: Arg #status predicted
F:42/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 5.2%; Score 119.0; DB 2; Length 423;
Best local similarity 22.9%; Pred. No. 11;
Matches 105; Conservative 46; Mismatches 188; Indels 120; Gaps 23;

QY 2 KAILIP--SLIPLTPSAFAQSEPELKLESWIVSRHGVRAPIKATQLMQDVTIPDAW 59
DB 10 QAAQLQFLQMLQWPPVQA-----RSLRFTLLYRHGDRSPVKA-----YPRDPYOE 58
QY 60 PTWPVKLGELTPGGELIAYLGHYWRQRLVADGLLPKCGCPOSGQVAIIADVD-ETRXT 118
DB 59 EKWPQGGQLTKGMLQHWELGQALRQY--HGFL--NASHYRQEVYVFSTDFDILMS 113
QY 119 GEAFAGLAGIADCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILLERAGGSTADPT 178
DB 114 AENLAGLPP-----TEQHNPNLSWQPI-----PVHTVPIE----- 148
QY 179 GHYQTAFLERLVNFPQSNICLKREKQDECSLTQALFSELKVASDCVSLTGAVSLAM 238
DB 149 -----DLIKFPLGHCPRYEQAGNE-----TRQIPYQMSIQNAQFLDMVANEIC 194
QY 239 LTELFL-----LQAQGMPEPGWGRITDSHOWNTLLSLHNAQF-----DLLQ 280
DB 195 LNNLILETWNVYDTLECRTHGLLEPPW---ASPTQVQALSQLKDFSLFLRGIHQDVQ 251
QY 291 PTFVAPSPAPVLDILKTALEPHDPQKQAYGVTVIPSVIFLACHDTNLANIGGALINW 340
DB 252 K-----APLQAVIQAQIKNIITIMATPSQF-----PKLIVYSAHDTTLVALQMLNVY- 300
QY 341 TLPGQDPTPGGELVFERRRRLSDNSQWLVQVSLVFLQCMKDKTELSINTPGEVKLI 400
DB 301 -----NKGAPVASHIFPIVQ--KDNVNF-SVMYFNDKSK---KAPW-----PLT 341
QY 401 LAGTEERNAQGMCSLAGTQIVNEARIP-----ACSLRS 434
DB 342 LPGGPR- ---CPQLQDFLRL-TEPVIPKDWKRECKLAS 374

RESULT 14
S64682
acid phosphatase (EC 3.1.3.2) Arph. i precursor, lysosomal - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 06 Dec 1996 #sequence revision 06-Dec-1996 #text change 17 Mar 2002
C:Accession: S64682; S64681
R:Chung, H.J.; Shaffer, C.; MacIntyre, R.
Mol. Gen. Genet. 250, 645-646, 1996
A:Title: Molecular characterization of the lysosomal acid phosphatase from Drosophila melanogaster
A:Reference number: S64681; MUID:96194627; PMID:8576866
A:Accession: S64682
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-438 <CHS>
A:Accession: S64681
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-31, 'Y', 33-438 <CHW>
C:Genetic codes:
A:Genes: FlyBase:FBan0000034
A:Cross-references: FlyBase:FBan0000034

C:Superfamily: mammalian acid phosphatase
C:Keywords: phosphoric monoester hydrolase
F:1-33/Domain: signal sequence #status predicted -SHM
F:34-436/Product: acid phosphatase Arph. i #status predicted -MAL

Query Match 4.9%; Score 112.5; DB 2; Length 488;
Best local similarity 20.6%; Pred. No. 0.48;
Matches 82; Conservative 61; Mismatches 147; Indels 109; Gaps 19;

QY 39 FLESWIVSRHGVRAPIKATQIMQVHTPDAMP-----WPVKLGELTPGGELIAYL 80
DB 47 GKRVHVIYFHSRTFV --- --LQVTFQWJGSEHTTQWCHFNLPQEHYDL 96
QY 81 GHYWRQRLVADGLLPKCGCPOSGQVAIIADVD-ETRXT 130
DB 96 GKWLPNRY--SNLLPPI---YSNENIYVQSTDVDTLMSAQSNLAGYFQGEDINWDI 150
QY 131 ---ATVITQADTSSPDPLFNPLKTGVCQLDNANVTDAILLERAGGSTADPT 187
DB 151 NWQPIPIHTSPERDP-----ILAAKAPCPAYDY-----E 180
QY 188 LERVNFPQSNICLKREKQDECSLTQALFSELKVASDCVSLTGAVSLAMTLTFLQ 247
DB 181 LASLESPEEK-ALTEKRNLFAYLSEKGRPVKTFID-----AQVLTNTLFLEN 229
QY 248 AGMPPEPGWGRITDSHOWNTLLSLHNAQFDLLQRTPEVARSRATPLL-DLIKIALTPHP 306
DB 230 LYNMILANIKKVVGREELTVYS--NFAFATSSYTKLAKLKAQPLVPTQRF----- 281
QY 307 QKQAYGVTLPTSVLFI-AGHDNTNLANGLALNWLPAQDNTFPGGELVFERRRRLSD 365
DB 282 KEKSSGSLKPKDRSMWVYSAHDTTVASVLNALK-FEL-----HSPYFACIMMELR--VD 434
QY 365 NSLWLVQVSLVFLQCMKDKTELSINTPGEVKLIAG 404
DB 334 ETNTEPLSVTFYK-----NTTAEPLPLDIPG 459

RESULT 15
S14742
acid phosphatase (EC 3.1.3.2) precursor - mouse (Mus musculus)
C:Species: Mus musculus (house mouse)
C:Date: 21 Nov 1993 #sequence revision 10-Nov-1993 #text change 21 Mar 2000
C:Accession: S14742
R:Geier, C.; von Figura, K.; Pohlmann, R.
Biochem. Biophys. Res. Commun. 172, 301-304, 1991
A:Title: Molecular cloning of the mouse lysosomal acid phosphatase
A:Reference number: S14742; MUID:91282986; PMID:2059437
A:Accession: S14742
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-421 <SEIF>
A:Cross-references: EMBL:X57199; NID:952979; PIDN:CAA00485.1; PID:952871
C:Superfamily: mammalian acid phosphatase
C:Keywords: lysosomo; phosphoric monoester hydrolase

Query Match 4.9%; Score 111.5; DB 2; Length 421;
Best local similarity 21.4%; Pred. No. 0.43;
Matches 98; Conservative 53; Mismatches 171; Indels 147; Gaps 24;

QY 2 KAILIP--LSLLIFTPSSAFASSEFTLKLESWIVSRHGVRAPIKATQIMQDVTIP 56
DB 8 QAAQLQFLQMLQWPPVQA-----RSLRFTLLYRHGDRSPVKA-----YPRDP 54
QY 57 DAMPTWPVKLGELTPGGELIAYLGHYWRQRLVADGLLPKCGCPOSGQVAIIADVD-ETR 115
DB 54 YQEEKWPQGGQLTKGMLQHWELGQALRQY--HGFL--NASHYRQEVYVFSTDFDRI 108
QY 116 RKTGEAFAGLAGIADCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILLERAG 175
DB 109 LMSAENLAGLPP-----NEVOHSPNLSWQPI-----PVHTVPIE----- 146
QY 175 GTTGHYQTAFLERLVNFPQSNICLKREKQDECSLTQALFSELKVASDCVSLTGAVSL 235

GenCore version 5.1.3
Copyright (c) 1993 - 2002 C. Murphy, et al

OM protein - protein search, using sw model

Run on: December 13, 2002, 15:49:35, Search time 40 Seconds
(without alignments)
1452.432 Million cell updates/sec

Title: US-09-866-379A-10

Perfect score: 2279

Sequence: 1 MKAILFFLELLPLPQEA.....GFTGIVNEAFIDNCSLESHL 436

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 14320626 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2300000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

| Database: A_Geneseq_101002.* | | | |
|------------------------------|--|--|--|
| 1: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1940.DAT* | | |
| 2: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1941.DAT* | | |
| 3: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1942.DAT* | | |
| 4: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1943.DAT* | | |
| 5: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1944.DAT* | | |
| 6: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1945.DAT* | | |
| 7: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1946.DAT* | | |
| 8: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1947.DAT* | | |
| 9: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1948.DAT* | | |
| 10: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1949.DAT* | | |
| 11: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1950.DAT* | | |
| 12: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1951.DAT* | | |
| 13: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1952.DAT* | | |
| 14: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1953.DAT* | | |
| 15: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1954.DAT* | | |
| 16: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1955.DAT* | | |
| 17: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1956.DAT* | | |
| 18: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1957.DAT* | | |
| 19: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1958.DAT* | | |
| 20: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1959.DAT* | | |
| 21: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1960.DAT* | | |
| 22: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1961.DAT* | | |
| 23: | /SILUS2/qcdat/a/geneseq/geneseq-emb1/AA1962.DAT* | | |

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 2288 | 97.8 | 430 | AAE15808 | Escherichia coli a |
| 2 | 2182 | 96.0 | 440 | AAE15809 | Escherichia coli a |
| 3 | 2188 | 96.0 | 440 | AAE15806 | Escherichia coli p |
| 4 | 2188 | 96.0 | 440 | AAE15806 | Escherichia coli B |
| 5 | 2182 | 95.7 | 432 | AAE15806 | Phytase associated |
| 6 | 2182 | 95.7 | 432 | AAE15806 | Lama2/APPA plasmid |
| 7 | 2182 | 95.7 | 432 | AAE15806 | R15/APPA plasmid |
| 8 | 2182 | 95.7 | 432 | AAE15806 | P15/APPA plasmid t |
| 9 | 2182 | 95.7 | 432 | AAE15806 | P15/APPA plasmid t |
| 10 | 2182 | 95.7 | 432 | AAE15806 | P15/APPA plasmid t |

| | | | | | |
|----|--------|------|------|----------|---------------------|
| 11 | 2182 | 95.7 | 442 | AAE15806 | SV40/APPA plasmid |
| 12 | 2182 | 95.7 | 442 | AAE15806 | Escherichia coli a |
| 13 | 2182 | 95.7 | 442 | AAE15806 | Escherichia coli a |
| 14 | 2177 | 95.5 | 432 | AAE15806 | E. coli acid phosph |
| 15 | 2171 | 95.3 | 432 | AAE15806 | E. coli acid phosph |
| 16 | 2170 | 95.2 | 432 | AAE15806 | E. coli acid phosph |
| 17 | 2157 | 94.6 | 432 | AAE15806 | E. coli acid phosph |
| 18 | 2147 | 94.2 | 432 | AAE15806 | E. coli acid phosph |
| 19 | 2080.5 | 91.3 | 423 | AAE15806 | AAE15806 |
| 20 | 1192 | 52.3 | 261 | AAE15806 | Novel human diurno |
| 21 | 637 | 28.0 | 144 | AAE15806 | Novel human diurno |
| 22 | 544 | 23.9 | 123 | AAE15806 | Novel human diurno |
| 23 | 418 | 13.6 | 118 | AAE15806 | Novel human diurno |
| 24 | 291 | 8.8 | 426 | AAE15806 | Novel human diurno |
| 25 | 291 | 8.8 | 426 | AAE15806 | Novel human diurno |
| 26 | 231 | 8.8 | 426 | AAE15806 | Novel human diurno |
| 27 | 184.5 | 8.1 | 1062 | AAE15806 | Novel human diurno |
| 28 | 131.5 | 5.8 | 423 | AAE15806 | Novel human diurno |
| 29 | 127 | 5.6 | 395 | AAE15806 | Novel human diurno |
| 30 | 126 | 5.5 | 440 | AAE15806 | Novel human diurno |
| 31 | 126 | 5.5 | 465 | AAE15806 | Novel human diurno |
| 32 | 126 | 5.5 | 455 | AAE15806 | Novel human diurno |
| 33 | 125 | 5.5 | 440 | AAE15806 | Novel human diurno |
| 34 | 125 | 5.5 | 440 | AAE15806 | Novel human diurno |
| 35 | 125 | 5.5 | 467 | AAE15806 | Novel human diurno |
| 36 | 124 | 5.4 | 467 | AAE15806 | Novel human diurno |
| 37 | 124 | 5.4 | 467 | AAE15806 | Novel human diurno |
| 38 | 124 | 5.4 | 467 | AAE15806 | Novel human diurno |
| 39 | 124 | 5.4 | 467 | AAE15806 | Novel human diurno |
| 40 | 123.5 | 5.4 | 392 | AAE15806 | Novel human diurno |
| 41 | 122 | 5.4 | 467 | AAE15806 | Novel human diurno |
| 42 | 118 | 5.2 | 440 | AAE15806 | Novel human diurno |
| 43 | 118 | 5.2 | 440 | AAE15806 | Novel human diurno |
| 44 | 117 | 5.1 | 582 | AAE15806 | Novel human diurno |
| 45 | 116.5 | 5.1 | 438 | AAE15806 | Novel human diurno |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| AAE15808 | |
| IL | AAE15806 Standard; Protein, 430 AA. |
| XX | |
| AC | AAE15808; |
| XX | |
| DT | 26 MAR 2002 (first entry) |
| XX | |
| DE | Escherichia coli appA phytase mutant protein. |
| XX | |
| XX | Bacterial Phytase; K12 appA phytase, phytase stability, diatolol; |
| KW | gastrointestinal, nutritional value, feed treatment process; therapy; |
| EW | thermal tolerance; growth performance; alcohol in drink; biopolymers; |
| KW | non-alcoholic drink; bioleaching; mutant; mutin. |
| XX | |
| OS | Escherichia coli. |
| CS | Synthetic |
| XX | |
| PH | Key Location/Qualifiers |
| FT | Misc-difference 68 /note= "Wild type Trp substituted with Glu" |
| FT | Misc-difference 84 /note= "Wild type Glu substituted with Trp" |
| FT | Misc-difference 95 /note= "Wild type Ala substituted with Pro" |
| FT | Misc-difference 97 /note= "Wild type Lys substituted with Cys" |
| FT | Misc-difference 168 /note= "Wild type Ser substituted with Glu" |
| FT | Misc-difference 180 /note= "Wild type Arg substituted with Tyr" |
| FT | Misc-difference 225 /note= "Wild type Asn substituted with Cys" |

cc nutritional value of phytate containing foodstuffs and subsequently
 cc improving the growth performance of an organism that consumes it, in
 cc treating animal digestive systems, in feed treatment processes and for
 cc in vitro purposes related to research, discovery and development; they
 cc are also used for generating recombinant digestive system like forms,
 cc for producing or manufacturing alcoholic and non-alcoholic drinks based
 cc on the use of moulds, grains and/or plants, in biorefining and bio-
 cc bleaching where a reduction in the use of environmentally harmful
 cc chemicals that are traditionally used in the pulp and paper industry
 cc is desired and in the reduction or possible elimination of the need
 cc for mineral supplements, enzymes or therapeutic drugs for animals
 cc from the daily feed thus increasing the amount calories and nutrients
 cc present in the feed. The present sequence is E. coli B phytase protein.

xx Sequence 440 AA;

Query Match 96.0%; Score 2188; DB 23; Length 440;
 Best Local Similarity 97.7%; Prod. No. 1.2c-212;
 Matches 425; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 MKATLIPFLSLIPITPSAFASQSEPEIKLESVVIVSPHGVAPATKATQLMQVTPDAMP 60
 DB 1 MKALLIPFLSLIPITPSAFASQSEPEIKLESVVIVSRHGVAPATKATQLMQDVTDPAMP 60
 QY 61 TWPVKLGPIITPGGELIAYLGHYWRQPIVAIGLLPKQGPQSSQVAITAIWDERTRKTGE 120
 DB 61 TWPVKLGPIITPGGELIAYLGHYWRQPIVAIGLLPKQGPQSSQVAITAIWDERTRKTGE 120
 QY 121 AFAGLAGAPKATIVHTQADISSPPDLPENPKLTGWCQIDNANVTDAILLRAGGSIADFTGH 180
 DB 121 AFAGLAGAPKATIVHTQADISSPPDLPENPKLTGWCQIDNANVTDAILLRAGGSIADFTGH 180
 QY 181 YQTAFFPEVINFPSQNIQIKPKQGPQSSQVAITAIWDERTRKTGE 240
 DB 181 ROTAFPEVINFPSQNIQIKPKQGPQSSQVAITAIWDERTRKTGE 240
 QY 241 EFLILQACQMPGPGWGRITDSHGNWTLISLHNAQFLLQETPEVARSRAPIPLDLIKTA 300
 DB 241 EFLILQACQMPGPGWGRITDSHGNWTLISLHNAQFLLQETPEVARSRAPIPLDLIKTA 300
 QY 301 LTPHPQKQAYGVTLPTSVLFIAQHDINANLGALEINWTLPGQDPTPPGGELVFERW 360
 DB 301 LTPHPQKQAYGVTLPTSVLFIAQHDINANLGALEINWTLPGQDPTPPGGELVFERW 360
 QY 461 KRSDNSQWTVQSVLFTQGMPEKPTLSINTEPPEVKLTLAGSEFNASQMGSLAGFTQ 420
 DB 461 KRSDNSQWTVQSVLFTQGMPEKPTLSINTEPPEVKLTLAGSEFNASQMGSLAGFTQ 420
 QY 421 TVNEARIPACSLRSH 445
 DB 421 TVNEARIPACSLRSH 445

RESULT 5
 AAB77775
 ID AAB77775 Standard; Protein: 442 AA.

xx AAB77775;
 xx 05-JUN-2002 (first entry)
 xx Phytase associated protein.
 xx Phytase.

xx Unidentified.
 xx KR99086028 A.

xx 15-DEC-1999.
 xx 25-MAY-1998; 98KR-0018810.

xx

PR 25-MAY-1998; 98KR-0018810.
 xx (W30AJ-0) W30AJN CO LTD.

xx Bao HD, Forcoburgh CW, Golden S, Chou KJ;

xx W01; 2000-045074/62.

xx N-PSDH; ABR12514.

xx Novel phytase gene, recombinant phytase and usage thereof

xx Disclosure; Fig 3; lopp; Korean.

cc The invention relates to a novel phytase gene, a recombinant
 cc phytase gene and their uses. This is the amino acid sequence of the
 cc phytase associated protein described in the invention.

xx Sequence 442 AA;

Query Match 95.7%; Score 2182; DB 21; Length 442;
 Best Local Similarity 98.1%; Prod. No. 4.8c-212;
 Matches 424; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MKATLIPFLSLIPITPSAFASQSEPEIKLESVVIVSPHGVAPATKATQLMQVTPDAMP 60
 DB 1 MKALLIPFLSLIPITPSAFASQSEPEIKLESVVIVSRHGVAPATKATQLMQDVTDPAMP 60
 QY 61 TWPVKLGPIITPGGELIAYLGHYWRQPIVAIGLLPKQGPQSSQVAITAIWDERTRKTGE 120
 DB 61 TWPVKLGPIITPGGELIAYLGHYWRQPIVAIGLLPKQGPQSSQVAITAIWDERTRKTGE 120
 QY 121 AFAGLAGAPKATIVHTQADISSPPDLPENPKLTGWCQIDNANVTDAILLRAGGSIADFTGH 180
 DB 121 AFAGLAGAPKATIVHTQADISSPPDLPENPKLTGWCQIDNANVTDAILLRAGGSIADFTGH 180
 QY 181 YQTAFFPEVINFPSQNIQIKPKQGPQSSQVAITAIWDERTRKTGE 240
 DB 181 ROTAFPEVINFPSQNIQIKPKQGPQSSQVAITAIWDERTRKTGE 240
 QY 241 EFLILQACQMPGPGWGRITDSHGNWTLISLHNAQFLLQETPEVARSRAPIPLDLIKTA 300
 DB 241 EFLILQACQMPGPGWGRITDSHGNWTLISLHNAQFLLQETPEVARSRAPIPLDLIKTA 300
 QY 301 LTPHPQKQAYGVTLPTSVLFIAQHDINANLGALEINWTLPGQDPTPPGGELVFERW 360
 DB 301 LTPHPQKQAYGVTLPTSVLFIAQHDINANLGALEINWTLPGQDPTPPGGELVFERW 360
 QY 461 KRSDNSQWTVQSVLFTQGMPEKPTLSINTEPPEVKLTLAGSEFNASQMGSLAGFTQ 420
 DB 461 KRSDNSQWTVQSVLFTQGMPEKPTLSINTEPPEVKLTLAGSEFNASQMGSLAGFTQ 420
 QY 421 TVNEARIPACSL 442
 DB 421 TVNEARIPACSL 442

RESULT 6
 AAB46257
 ID AAB46257 Standard; Protein: 442 AA.

xx AAB46257;

xx 20-FEB-2001 (first entry)

xx Lama2/Abpa plasmid translated sequences.

xx Transgenic animal; salivary protein; phytase; Phosphate-animat, animal

xx environmental pollution; pig.

xx Mus musculus - chimeric.

xx Escherichia coli - chimeric.

xx WO200064247-A1.


```

XX 02-NOV-2000.
PD
XX
XX 20 APR 2000; 2000WO-CA00430.
PF
XX
XX 24-APR-1999; 94US-0130508.
PR
XX (UYGU ) UNIV GUELPH.
PA
XX Forsberg CW, Golovan S, Phillips JP;
PI
XX WPI; 2000-687245/67.
DR
XX N-PSDB; AAC68294.
DR
XX
XX Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein -
XX
XX Disclosure; Fig 5; 152pp; English.
PS
XX
XX The present invention provides transgenic animals which produce desired
CC proteins in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence.
XX
XX Sequence 432 AA:
SQ
Query Match 95.7%; Score 2182; DB 21; Length 432;
Best Local Similarity 98.1%; Pred. No. 4.80-212;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0.
QY 1 MKAILLPILLSLIPITPQSAFAQSPPELKLSVIVSRHCVAPTKATQIMQVTPDAMP 60
DB 1 MKAILLPILLSLIPITPQSAFAQSEPELKLSVIVSRHCVAPTKATQIMQVTPDAMP 60
QY 61 TWPVKLGWLTTPGGRIIAYIGHYQKQPIVAGCIIPKQGPQSQVAILAIWVPEPKTGE 120
DB 61 TWPVKLGWLTTPGGRIIAYIGHYQKQPIVAGCIIPKQGPQSQVAILAIWVPEPKTGE 120
QY 121 AFAAGLAPGCAIVHTQADTSSPDPLNPKTKGVQOLDNANVTDAILERAGSIAFTGH 180
DB 121 AFAAGLAPGCAIVHTQADTSSPDPLNPKTKGVQOLDNANVTDAILERAGSIAFTGH 180
QY 181 YOTAFPELEFVINFQSNLCLEPKKQESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 YOTAFPELEFVINFQSNLCLEPKKQESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EFLAQAGAMPPEQWGRITTSQWNNTILSLHNAQFYLLQPTPEVAPSPATPLDLIKTA 300
DB 241 EFLAQAGAMPPEQWGRITTSQWNNTILSLHNAQFYLLQPTPEVAPSPATPLDLIKTA 300
QY 301 LTPHPQKQAYGVTLPTSVLTAGHDNLANLGAGLELNWTLPGQPDNTPPGSELVERW 360
DB 301 LTPHPQKQAYGVTLPTSVLTAGHDNLANLGAGLELNWTLPGQPDNTPPGSELVERW 360
QY 361 PLSNSQWQVSIQVSLVFTQIQMPKPKLSINTPGEVKTLTACCFERNAGCMCSLAGTQ 420
DB 361 PLSNSQWQVSIQVSLVFTQIQMPKPKLSINTPGEVKTLTACCFERNAGCMCSLAGTQ 420
QY 421 LVNEARIPACSL 432
DB 421 LVNEARIPACSL 432
RESULT 7
AAB36258
ID AAB46258 standard; Protein; 432 AA
XX
AC AAB46258;
XX

```

```

DT 20-FEB-2001 (first entry)
XX
XX R15/APPA plasmid translated sequence.
XX
XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW environmental pollution; pig.
XX
XX Rattus sp - chimeric.
OS Escherichia coli - chimeric.
XX
XX WO2000064247-A1.
PN
XX 02-NOV-2000.
PD
XX
XX 20-APR-2000; 2000WO-CA00430.
PF
XX 24-APR-1999; 94US-0130508.
PR
XX (UYGU-) UNIV GUELPH.
PA
XX Forsberg CW, Golovan S, Phillips JP;
PI
XX WPI; 2000-687245/67.
DR
XX N-PSDB; AAC68295.
DR
XX
XX Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein -
XX
XX Disclosure; Fig 18; 152pp; English.
PS
XX
XX The present invention provides transgenic animals which produce desired
CC proteins in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence.
XX
XX Sequence 432 AA:
SQ
Query Match 95.7%; Score 2182; DB 21; Length 432;
Best Local Similarity 98.1%; Pred. No. 4.80-212;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0.
QY 1 MKAILLPILLSLIPITPQSAFAQSPPELKLSVIVSRHCVAPTKATQIMQVTPDAMP 60
DB 1 MKAILLPILLSLIPITPQSAFAQSEPELKLSVIVSRHCVAPTKATQIMQVTPDAMP 60
QY 61 TWPVKLGWLTTPGGRIIAYIGHYQKQPIVAGCIIPKQGPQSQVAILAIWVPEPKTGE 120
DB 61 TWPVKLGWLTTPGGRIIAYIGHYQKQPIVAGCIIPKQGPQSQVAILAIWVPEPKTGE 120
QY 121 AFAAGLAPGCAIVHTQADTSSPDPLNPKTKGVQOLDNANVTDAILERAGSIAFTGH 180
DB 121 AFAAGLAPGCAIVHTQADTSSPDPLNPKTKGVQOLDNANVTDAILERAGSIAFTGH 180
QY 181 YOTAFPELEFVINFQSNLCLEPKKQESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 YOTAFPELEFVINFQSNLCLEPKKQESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EFLAQAGAMPPEQWGRITTSQWNNTILSLHNAQFYLLQPTPEVAPSPATPLDLIKTA 300
DB 241 EFLAQAGAMPPEQWGRITTSQWNNTILSLHNAQFYLLQPTPEVAPSPATPLDLIKTA 300
QY 301 LTPHPQKQAYGVTLPTSVLTAGHDNLANLGAGLELNWTLPGQPDNTPPGSELVERW 360
DB 301 LTPHPQKQAYGVTLPTSVLTAGHDNLANLGAGLELNWTLPGQPDNTPPGSELVERW 360
QY 361 PLSNSQWQVSIQVSLVFTQIQMPKPKLSINTPGEVKTLTACCFERNAGCMCSLAGTQ 420
DB 361 PLSNSQWQVSIQVSLVFTQIQMPKPKLSINTPGEVKTLTACCFERNAGCMCSLAGTQ 420

```

QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432

RESULT 8
 AAB36259
 ID AAB36259 standard; Protein; 432 AA.
 AC AAB36259;
 XX
 DT 20-FEB-2001 (first entry)
 DE R15/APPa plasmid translated sequence
 XX
 KW Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 OS Rattus sp - chimeric.
 XX Escherichia coli - chimeric.
 XX
 DN W0200064247-A1.
 XX
 FD 02-NOV-2000.
 XX
 PF 20-APR-2000; 2000WO-CA00430.
 XX
 PR 23-APR-1999; 990S-0140508.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX
 PI Forsberg CW, Golovan S, Phillips JP;
 XX
 DR WPI; 2000-687245/67.
 DR N-PSDB; AAC68296.
 XX
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein
 XX
 PS Disclosure; Fig 19; 152pp; English.
 XX
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which express phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPa phytase coding sequence.
 XX
 SQ Sequence 432 AA;

Query Match 66.7%; Score 2182; DB 21; Length 432;
 Best Local Similarity 98.1%; Pred. No. 4, 8e-212;
 Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0.

QY 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSVSHGVRAPTKATQLMQDVTTPDAMP 60
 DB 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSVSHGVRAPTKATQLMQDVTTPDAMP 60
 QY 61 TWPVKLGELTPRGGFIAYIGHYWGPQRTVAIGLIGPKGTGPGSGGVALIAPVDEPTKGTGE 120
 DB 61 TWPVKLGELTPRGGFIAYIGHYWGPQRTVAIGLIGPKGTGPGSGGVALIAPVDEPTKGTGE 120
 QY 121 APAAGIAPVCAITVHTQADTSPPDIPNLKFTVYVQLINANVTDAITEPAGCSSTADPTGH 180
 DB 121 AFAAGIAPVCAITVHTQADTSPPDIPNLKFTVYVQLINANVTDAITEPAGCSSTADPTGH 180
 QY 181 YQTAFPIFRVNIPTQSNLCIKPKQDESLSITQALPSSELKVSADCVSLTQAVSLASMLT 240
 DB 181 YQTAFPIFRVNIPTQSNLCIKPKQDESLSITQALPSSELKVSADCVSLTQAVSLASMLT 240
 QY 241 FIFLIQQAQMPGPCWGRITDSHQWNTLILSHNAQFYLLQRTPEVARSPATPLDLIKTA 300

DB 241 FIFLIQQAQMPGPCWGRITDSHQWNTLILSHNAQFYLLQRTPEVARSPATPLDLIKTA 300
 QY 301 ITPHPQQAQYGVTLPTSVLFTAGHDTNLANAGGALLEHWLFLSGHRTTPESFVFEW 400
 DB 301 ITPHPQQAQYGVTLPTSVLFTAGHDTNLANAGGALLEHWLFLSGHRTTPESFVFEW 400
 QY 361 KKLISNSQWLVQSVLVLQGMKDKLPSLNTGCEVALALAVGELTNAGSIMESSAGFTQ 420
 DB 361 KKLISNSQWLVQSVLVLQGMKDKLPSLNTGCEVALALAVGELTNAGSIMESSAGFTQ 420
 QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432

RESULT 9
 AAB36260
 ID AAB36260 standard; Protein; 432 AA.
 XX
 AC AAB36260;
 XX
 DT 20-FEB-2001 (first entry)
 DE R15/APPa plasmid translated sequence.
 XX
 KW Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 OS Rattus sp - chimeric.
 XX Escherichia coli - chimeric.
 XX
 DN W0200064247-A1.
 XX
 FD 02-NOV-2000.
 XX
 PF 20-APR-2000; 2000WO-CA00430.
 XX
 PR 23-APR-1999; 990S-0140508.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX
 PI Forsberg CW, Golovan S, Phillips JP;
 XX
 DR WPI; 2000-687245/67.
 DR N-PSDB; AAC68297.
 XX
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein
 XX
 PS Disclosure; Fig 20; 152pp; English.
 XX
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which express phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPa phytase coding sequence.
 XX
 SQ Sequence 432 AA;

Query Match 95.7%; Score 2182; DB 21; Length 432;
 Best Local Similarity 98.1%; Pred. No. 4, 8e-212;
 Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0.

QY 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSVSHGVRAPTKATQLMQDVTTPDAMP 60
 DB 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSVSHGVRAPTKATQLMQDVTTPDAMP 60
 QY 61 TWPVKLGELTPRGGFIAYIGHYWGPQRTVAIGLIGPKGTGPGSGGVALIAPVDEPTKGTGE 120

| | | | |
|-----|----|---|-----|
| 61 | DB | TRIVFVLCNMLPEQZDEI IAVI RHYQZQBI VAIQEL AKKQCPQSUGVA I IADVHERTKTGE | 120 |
| 121 | QY | AFAAGLAPDCAITVHTQAITSSDPDLFNLPLTKGVCOLONANTDAILLRAGSSTADFTGH | 180 |
| | | | |
| 121 | DB | AFAAGLAPDCAITVHTQAITSSDPDLFNLPLTKGVCOLONANTDAILLRAGSSTADFTGH | 180 |
| | | | |
| 181 | QY | YQTAPELEPEVLNFPQSNLCLEKPEFKQFSSLTQALPSELKVSADVCVSTIGAVSASMLT | 240 |
| | | | |
| 181 | DB | YQTAPELEPEVLNFPQSNLCLEKPEFKQFSSLTQALPSELKVSADVCVSTIGAVSASMLT | 240 |
| | | | |
| 241 | QY | EIEELLQQAQCMPEPWCRITTDISHQWNTLISIHNNQFDILQPTPEVARSPPATPLIDLIKTA | 300 |
| | | | |
| 241 | DB | EIEELLQQAQCMPEPWCRITTDISHQWNTLISIHNNQFDILQPTPEVARSPPATPLIDLIKTA | 300 |
| | | | |
| 301 | QY | LTPHPQCKQAYGYVLTSLVLTAGHDNTIANIAGALELAWTLPGPDNTPPGGELVFEFW | 360 |
| | | | |
| 301 | DB | LTPHPQCKQAYGYVLTSLVLTAGHDNTIANIAGALELAWTLPGPDNTPPGGELVFEFW | 360 |
| | | | |
| 361 | QY | FSYSSNSWLVQVSI VEQTICQMPKDTPI SINTPPQEVKTLTACGFEPNACGMC'SLACGTTQ | 420 |
| | | | |
| 361 | DB | FSYSSNSWLVQVSI VEQTICQMPKDTPI SINTPPQEVKTLTACGFEPNACGMC'SLACGTTQ | 420 |
| | | | |
| 421 | QY | LVNPARIPACSL 432 | |
| | | | |
| 421 | DB | LVNPARIPACSL 432 | |

RESULT 10
 AAR36261
 10 AAR36261 standard; Protein; 432 AA.
 XX
 AC AAR36261;
 XX
 PT 20-FEB-2001 (first entry)
 XX
 DE R15/Alpha plasmid translated sequence.

| | |
|----|--|
| KW | Transgenic animal; salivary protein; phytase; pneumonitis; animal growth |
| KW | environmental pollution; pig. |
| XX | |
| XX | Rattus sp - chimeric. |
| OS | Escherichia coli - chimeric. |
| XX | |
| XX | WO2000064247-A1. |
| PM | |
| XX | |
| XX | 02-NOV-2000. |
| XX | |
| XX | 29-APR-2000; 2000WO-C000430 |
| XX | |
| XX | 23-APR-1999; 990S-0130508 |
| XX | |
| XX | (OYGD-) UNIV GUELPH. |
| XX | |
| PI | Forsberg CW, Golovan S, Phillips JP; |
| XX | |
| XX | WPI: 2000-687245/67. |
| DR | N-PS08; AAC:68298. |
| XX | |
| XX | Transgenic non-human animal for gastrointestinal tract specific |
| PT | expression of a protein, preferably phytase, comprises a nucleic acid |
| PT | sequence including a heterologous transgene construct encoding the |
| PT | protein - |
| PI | |
| XX | |
| XX | Disclosure; Fig 21: 152pp; English. |
| PS | |
| XX | |
| CC | The present invention provides transgenic animals which produce desired |
| CC | proteins, in this case pigs which expresses phytase in the salivary |
| CC | gland, low phytase production levels result in phytate in the diet being |
| CC | excreted and causing phosphorus contamination in water, as well as |
| CC | reducing the growth of animals. The invention provides a number of |
| CC | transgenes containing the E. coli APFA phytase coding sequence. |
| XX | |
| XX | Sequence 432 AA; |
| SQ | |

| Query Match | 95.7% | Score 2182 | EB 21 | Length 432 |
|-----------------------|--------------|---|--------------|------------|
| Best Local Similarity | 98.1% | Pred. No. 4.8e-212 | | |
| Matches 424 | Conservative | 0 | Mismatches 8 | Indels 0 |
| QY | 1 | KKALLPFLSLLPLTPQSAFAQSEPRKLFESVVIYSRHGVRAPTKAIQIMQVIVIPAWP | 60 | |
| DB | 1 | KKALLPFLSLLPLTPQSAFAQSEPRKLFESVVIYSRHGVRAPTKAIQIMQVIVIPAWP | 60 | |
| QY | 61 | TPWPKUGELTPPGGELLAYLGHYWKQVLKVALGILLPKAGGQSGQVAVIAAVDERTKTE | 120 | |
| DB | 61 | TPWPKUGELTPPGGELLAYLGHYWKQVLKVALGILLPKAGGQSGQVAVIAAVDERTKTE | 120 | |
| QY | 121 | AFAGLAPDCAITVHTQADTSSPDLPFNPKIGVQQLDANVANVDAILEKAGSGIADFTGH | 180 | |
| DB | 121 | AFAGLAPDCAITVHTQADTSSPDLPFNPKIGVQQLDANVANVDAILEKAGSGIADFTGH | 180 | |
| QY | 181 | YQTAPELEPRVINFPPQSNGLCKPEKQDESLSLTQALPSELKVSADCVSLIGAVSLASMLI | 240 | |
| DB | 181 | YQTAPELEPRVINFPPQSNGLCKPEKQDESLSLTQALPSELKVSADCVSLIGAVSLASMLI | 240 | |
| QY | 241 | ELELLEACAMPPEKWKILFDSHOWNILLSEHNAFPLELEPRFVAVSFATLLEPLFA | 300 | |
| DB | 241 | ELELLEACAMPPEKWKILFDSHOWNILLSEHNAFPLELEPRFVAVSFATLLEPLFA | 300 | |
| QY | 301 | LTPHPKQKQAVGVTLPSTSVLF IAGHDINLANIGALELNLWTLPGQPDTPPGELVFERW | 360 | |
| DB | 301 | LTPHPKQKQAVGVTLPSTSVLF IAGHDINLANIGALELNLWTLPGQPDTPPGELVFERW | 360 | |
| QY | 361 | PRLSDNSQWIGVSLVFOTLQOMRKTPILSLNPPGFVKLLIAGSEPRNAQMCNLSAGTQ | 420 | |
| DB | 361 | PRLSDNSQWIGVSLVFOTLQOMRKTPILSLNPPGFVKLLIAGSEPRNAQMCNLSAGTQ | 420 | |
| QY | 421 | IVNEARIPACSL 432 | | |
| DB | 421 | IVNEARIPACSL 432 | | |

RESULT 11
 AAB36262
 ID AAB36262 standard; Protein: 432 AA.
 XX AC
 XX AAB36262;
 XX DT 20-FEB-2001 (first entry)
 XX DE SV40/APFA plasmid translated sequence.
 XX XX
 XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 XX XX
 XX Rhesus macaque polyoma virus - chimeric.
 OS Escherichia coli - chimeric.
 XX XX
 XX WO2000064247-A1.
 XX XX
 XX 02-NOV-2000.
 XX XX
 XX 20-APR-2000; 2000WO-CA00430.
 XX XX
 XX 23-APR-1999; 99US-0130508.
 XX XX
 XX (UYGU-) UNIV GUELPH.
 XX PA
 XX Forsberg CW, Golovan S, Phillips JP;
 PT PI
 XX WPI; 2000 687245/67.
 XX CR N-PSDB; AAC68299.
 XX DR
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein -

| XX | disclosure: Fig 22; 152bp; English. |
|----|--|
| XX | The present invent ion provides transgenic animals which produce desired |
| XX | proteins, in this case, phytase, which express a phytase in the phytase |
| XX | gland. Low phytase production levels result in phytate in the diet being |
| XX | excreted and causing phosphorus contamination in water, as well as |
| XX | reducing the growth of animals. The invention provides a number of |
| XX | transgenes containing the E. coli APHA phytase coding sequence. |
| XX | Sequence 432 AA: |
| XX | Query Match 95.7%; Eseq 2192; PB 21; Length 432; |
| XX | Best local Similarity 98.1%; Prod. No. 4,80-212; |
| XX | Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0; |
| QY | 1 MKATLTPFSLTTPQSAFAFAGSEPFPIKIFGVVIVSPHCVPAPATATATMCGVTPAWP 60 |
| DB | 1 MKALLPELSLAPITPQSAFAFAGSEPELKLSVIVSRHGVAPRATKATQLMODVTPAWP 60 |
| QY | 61 TWEVKLGFTPRGSEIAYIGHYWPQPIVADQII PKGSPQSCQVALIAVDDEPTPTGE 120 |
| DB | 61 TWEVKLGFTPRGSEIAYIGHYWPQPIVADQII PKGSPQSCQVALIAVDDEPTPTGE 120 |
| QY | 61 TWEVKLGFTPRGSEIAYIGHYWPQPIVADQII PKGSPQSCQVALIAVDDEPTPTGE 120 |
| DB | 61 TWEVKLGFTPRGSEIAYIGHYWPQPIVADQII PKGSPQSCQVALIAVDDEPTPTGE 120 |
| QY | 121 AFAAGLAPICATIVHTQADTSSPDPIFNPIKTCWCOLDANVNTDAILERAGGSLADFTGH 180 |
| DB | 121 AFAAGLAPICATIVHTQADTSSPDPIFNPIKTCWCOLDANVNTDAILERAGGSLADFTGH 180 |
| QY | 181 YGTAFRETFVLNFPQSNCLKREKQDESSLTQALFSELKVSADNVSLTCGVASLASMLT 240 |
| DB | 181 RQTAFRELEKVLNFPQSNCLKREKQDESSLTQALFSELKVSADNVSLTCGVASLASMLT 240 |
| QY | 241 EFLDAQACMPFGWGRITDSQWNTLLSLHNAQFDLQRTPEVARSRATPLDLDKTA 300 |
| DB | 241 EFLDAQACMPFGWGRITDSQWNTLLSLHNAQFDLQRTPEVARSRATPLDLDKTA 300 |
| QY | 301 LTPHPQKQAYGVTHPTSVLFIACHDNTLANLGGALELNWTLPGQPIPTPGGELVFERW 360 |
| DB | 301 LTPHPQKQAYGVTHPTSVLFIACHDNTLANLGGALELNWTLPGQPIPTPGGELVFERW 360 |
| QY | 361 RRLSUNSQWQLQVSLVFQTLQGMKTRPISINTPPEVKITLAVCFEPNAGCM-SLAGE-TQ 420 |
| DB | 361 RRLSUNSQWQLQVSLVFQTLQGMKTRPISINTPPEVKITLAVCFEPNAGCM-SLAGE-TQ 420 |
| QY | 421 TVNEARIPACSL 432 |
| DB | 421 TVNEARIPACSL 432 |

| | |
|-----------|--|
| RESULT 12 | |
| AA036263 | |
| ID | AA036263 standard; Protein; 442 AA. |
| XX | |
| AC | AA036263; |
| XX | |
| PI | 20 FEB 2001 (first entry) |
| XX | |
| DE | Lam42/ATTA plasmid translated sequence. |
| XX | |
| KW | transgenic animal; salivary protein; phytase; phosphorus; animal growth; |
| KW | environmental pollution; pig. |
| XX | |
| OS | Mus musculus chimeric. |
| OS | Escherichia coli - chimeric. |
| XX | |
| PN | WC2000064247-A1. |

| | | |
|----|-----|---|
| PA | XX | (UUGD-) UNIV GDELPH. |
| XX | XX | |
| PI | XX | Forsberg CW, Golovan S, Phillips D; |
| XX | XX | |
| XX | XX | W22: 273-274; 275-277. |
| XX | XX | |
| DR | XX | N PSUK; AA068400. |
| XX | XX | |
| XX | XX | Transgenic non-human animal for gastrointestinal tract specific |
| PT | PT | expression of a protein, preferably phytase, comprising a nucleotide acid |
| PT | PT | sequence including a heterologous transgene construct encoding the |
| PT | PT | protein. |
| XX | XX | |
| XX | XX | Disclosure: Fig 24; 152pp; English. |
| XX | XX | |
| CC | CC | The present invention provides transgenic animals which produce desirable |
| CC | CC | proteins. In this case pigs which expresses phytase in the salivary |
| CC | CC | gland, low phytase production levels result in phytate in the diet being |
| CC | CC | excreted and causing phosphorus contamination in water, as well as |
| CC | CC | reducing the growth of animals. The invention provides a number of |
| XX | XX | transgenes containing the E. coli APPA phytase coding sequence. |
| XX | XX | |
| XX | XX | Sequence 432 AA: |
| XX | XX | |
| XX | XX | Query Match 95.7%; Score 2182; DB 21; Length 432; |
| XX | XX | Best Local Similarity 98.1%; Pred. No. 4,86 212; |
| XX | XX | Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0 |
| QY | 1 | MAAILPTELILPLTPQSAFAQSEPELKLESVIVSRHVPATPKALQMGVIVAWP 60 |
| DB | 1 | MAAILPTELILPLTPQSAFAQSEPELKLESVIVSRHVPATPKALQMGVIVAWP 60 |
| QY | 61 | TPWVKLGLTPRGELIAYLGHWYQRIYALGKPKYQPSQAVATIAIWIPEKKE 120 |
| DB | 61 | TPWVKLGLTPRGELIAYLGHWYQRIYALGKPKYQPSQAVATIAIWIPEKKE 120 |
| QY | 121 | AFAAGLAPICATIVHQAITSIDPLENLKIGVQLINANVHAIILFAVGSIAIDTGH 180 |
| DB | 121 | AFAAGLAPICATIVHQAITSIDPLENLKIGVQLINANVHAIILFAVGSIAIDTGH 180 |
| QY | 181 | YOTAFETI ERVINFPQSNICTKREKQDSQSTQALPSELKVSALPVSLLIYVSLASMLI 240 |
| DB | 181 | ROTAFETI ERVINFPQSNICTKREKQDSQSTQALPSELKVSALPVSLLIYVSLASMLI 240 |
| QY | 241 | EPLTQQAGSMPEPCWGRITTSIDWNHLLSLHNAQFIDILQTFEYAVKSPATLIDILKIA 300 |
| DB | 241 | EPLTQQAGSMPEPCWGRITTSIDWNHLLSLHNAQFIDILQTFEYAVKSPATLIDILKIA 300 |
| QY | 301 | LTHHPQKQAYCVTLPTTSVLFITAGHDIFIANIGYALFINWII PQGIDNITPDSGVIFPEW 360 |
| DB | 301 | LTHHPQKQAYCVTLPTTSVLFITAGHDIFIANIGYALFINWII PQGIDNITPDSGVIFPEW 360 |
| QY | 361 | RRIISNSIATQVSLVFPDGLQMPKPTPSLNTIPGQVWLLIAGVFEENALQMSIAIHL 420 |
| DB | 361 | RRIISNSIATQVSLVFPDGLQMPKPTPSLNTIPGQVWLLIAGVFEENALQMSIAIHL 420 |
| QY | 421 | IVNEARIPACSL 432 |
| DB | 421 | IVNEARIPACSL 432 |

| | |
|-----------|--|
| RESULT 13 | |
| AAE15807 | |
| ID | AAE15807 standard; Protein; 442 AA. |
| XX | |
| AC | AAE15807; |
| XX | |
| DT | 26-MAR-2002 (first entry) |
| XX | |
| DE | <i>Escherichia coli</i> appA phytase wild type protein.. |
| XX | |
| KW | Bacterial phytase; K12 appA phytase; protease stability; anabolism; gastrointestinal; nutritional value; food treatment process; therapy; thermal tolerance; growth performance; alcoholic drink; biodegradation |

KW non-alcoholic drink; biobleaching.
 XX
 OS Escherichia coli.
 PT
 PN W0200190333-A2.
 PP
 XX 29-NOV-2001.
 XX
 XX 24-MAY-2001; 2001WO-US17118.
 XX
 XX 25-MAY-2000; 2000US-0580515.
 PR
 XX (DIVE-) DIVERSA CORP.
 PA
 XX Short JM, Kretz KA, Gray KA, Barton NP, Garrett JH, O'Donoghue E;
 PT
 DR WPI; 2002-083108/11.
 DR N-PSDB; AA025463.
 XX
 XX New bacterial phytase for e.g. improving the nutritional value of
 PT phytate-containing foodstuffs and subsequently improving the growth
 PT performance of an organism that consumes it, or in treating animal
 PT digestive systems
 XX
 XX claim 54; Fig 8; 170pp; English.
 PS
 XX The patent discloses recombinant bacterial phytase from Escherichia coli
 CC E.2.4p4 phytase. The enzyme has phytase activity and improved thermal
 CC tolerance when compared with wild type phytase. It has improved protease
 CC stability at low pH. The recombinant phytase is useful for improving the
 CC nutritional value of phytate-containing foodstuffs and subsequently
 CC improving the growth performance of an organism that consumes it, in
 CC treating animal digestive systems, in feed treatment processes and for
 CC in vitro purposes related to research, discovery and development. They
 CC are also used for generating recombinant digestive system like forms,
 CC for producing or manufacturing alcoholic and non-alcoholic drinks based
 CC on the use of moulds, grains and/or plants, in biopulping and bio-
 CC bleaching where a reduction in the use of environmentally harmful
 CC chemicals that are traditionally used in the pulp and paper industry
 CC is desired and in the reduction or possible elimination of the need
 CC for mineral supplements, enzymes or therapeutic drugs for animals
 CC from the daily feed thus increasing the amount calories and nutrients
 CC present in the feed. The present sequence is E. coli appA phytase
 CC wild type protein.
 XX
 SQ Sequence 432 AA;

Query Match 95.7%, Score 2192; DB 23; Length 432.
 Best Local Similarity 98.1%; Pred. No. 4, 8c-212;
 Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKALLPFLSLILPLTQSAFAQSEPELKLESVVVSRHGVRAPTKATQIMQVILDAWF 60
 DB 1 MKALLPFLSLILPLTQSAFAQSEPELKLESVVVSRHGVRAPTKATQIMQVILDAWF 60
 QY 61 TWPVKLGLTPGGLTAYLGHYQRLVADGKLLAKKQCPQSGVALLADVDETRKTGE 120
 DB 61 TWPVKLGLTPGGLTAYLGHYQRLVADGKLLAKKQCPQSGVALLADVDETRKTGE 120
 QY 121 AFAANAPCAATVHTAGVSCRPENPKEVYQIDNANVTIATPEAGGSLADPTGH 180
 DB 121 AFAAGLAPCAATVHTAGVSCRPENPKEVYQIDNANVTIATPEAGGSLADPTGH 180
 QY 181 YCTAPPELPVILNPSQNCIKPKQDPSCSTQATPSEIKVSDCVSLTGAVSLASMLT 240
 DB 181 KCTAPPELPVILNPSQNCIKPKQDPSCSTQATPSEIKVSDCVSLTGAVSLASMLT 240
 QY 241 EFLTQQAGQMPPEPCWCPITTSQWNTTILSHNAQFDLLOPTPEVAVSRATPLDLTKTA 300
 DB 241 EFLTQQAGQMPPEPCWCPITTSQWNTTILSHNAQFDLLOPTPEVAVSRATPLDLTKTA 300
 QY 401 LTPHPPOKQAVGVTLPTSVLFIAGHDNLANLGGALFLNWTLLNQDNTPTPGGLVFERW 360
 DB 401 LTPHPPOKQAVGVTLPTSVLFIAGHDNLANLGGALFLNWTLLNQDNTPTPGGLVFERW 360

DB 301 LTPHPPOKQAVGVTLPTSVLFIAGHDNLANLGGALFLNWTLLNQDNTPTPGGLVFERW 460
 QY 361 RRLSDNSQWITQVSLVFTLQOMKDKTFLSLNTPPEVKELIAGCEERNAQWPSLAFTQ 420
 DB 361 RRLSDNSQWITQVSLVFTLQOMKDKTFLSLNTPPEVKELIAGCEERNAQWPSLAFTQ 420
 QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432
 RESULT 14
 ID AAE02631 standard, Protein; 432 AA.
 XX AAE02631:
 XX 06-AUG-2001 (first entry)
 XX E. coli acid phosphatase/phytase (appa) protein.
 DE Acid phosphatase-phytase; appA, enzymatic activity; food additive;
 XX animal feed; monoester phosphate; cereal food; human food.
 KW Escherichia coli.
 XX
 XX Key location/qualifiers
 FH Misc-difference 109 /note= "Encoded by GPC"
 FT Misc-difference 148 /note= "Encoded by ATT"
 FT Modified-site 161..163 /note= "Asn is N-glycosylated"
 FT Disulfide-bond 200..210
 FT Region 202..211 /label= GH_loop
 FT Modified-site 339..341 /note= "Asn is N-glycosylated"
 XX W0200136607-A1
 XX 25-MAY-2001.
 XX 17-NOV-2000; 2000WO-US31622.
 XX 18-NOV-1999; 99US-0166179.
 XX (COPP) CORNELL RES FOUND INC.
 XX Lei X;
 XX WPI; 2001-367579/98
 XX N-PSDB; AA06831.
 PT Mutated acid phosphatase/phytase from Escherichia coli has improved
 PT enzymatic activity compared to the wild type and is useful as a food
 PT additive, particularly for animal feeds
 XX
 PS Claim 1; Page 11-12; 56pp; English.
 XX The present sequence is wild type acid phosphatase/phytase (PPA)
 CC from Escherichia coli. The wild type appA is mutated to enhance
 CC its enzymatic activity. Phytases, a specific group of monoester
 CC phosphatases, are required to initiate the release of phosphate from
 CC phytate, the major storage of phosphate in cereal foods or feeds. The
 CC mutant acid phosphatase/phytase is added to animal feed as a food
 CC additive to improve uptake of phosphate and zinc from the diet. The
 CC phytase can also be added to human food. The mutant phytase is more heat
 CC stable than phytase isolated from Aspergillus niger and safer for use
 CC in human food manufacture
 XX
 SQ Sequence 432 AA;

Query Match 95.5% Score 2177; DP 22; Length 432;
 Best Local Similarity 97.7%; Pred. No. 6,26 211;
 Matches 422; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKATLIPFSLIIPITPQSAFASQSFPIKLESVVIVSRHCVAPTRATQIMQVITPAMP 60
 DB 1 MKATLIPFSLIIPITPQSAFASQSFPIKLESVVIVSRHCVAPTRATQIMQVITPAMP 60

QY 61 TWVVKIGFTTPRGFTIAYIGHYWQRIYVAQHIKPGGTCQSCQVAIIADVDPSTKTF 120
 DB 61 TWVVKIGFTTPRGFTIAYIGHYWQRIYVAQHIKPGGTCQSCQVAIIADVDPSTKTF 120

QY 121 AFAGLAPICATIVHTGATSSPHFENLKTGWAQDENANVTGAILERAGSLADEF 180
 DB 121 AFAGLAPICATIVHTGATSSPHFENLKTGWAQDENANVTGAILERAGSLADEF 180

QY 181 YQAFRELEVLPQSNLCIEPEKQESQSLTQALPSELKVSADCVSLGAVSLASMLT 240
 DB 181 YQAFRELEVLPQSNLCIEPEKQESQSLTQALPSELKVSADCVSLGAVSLASMLT 240

QY 241 ETELQQAQCMPEGWGRITDSHOWNTLSIHNAQPHLQRTPEVAKSRATPLILIKTA 300
 DB 241 ETELQQAQCMPEGWGRITDSHOWNTLSIHNAQPHLQRTPEVAKSRATPLILIKTA 300

QY 301 LIDPHPPKQAYGVITPISVLEIACHDINLANIGALELNWITLQAPENITFVEIPEEW 360
 DB 301 LIDPHPPKQAYGVITPISVLEIACHDINLANIGALELNWITLQAPENITFVEIPEEW 360

QY 361 RRLSDNSQWLOVSLVFTQIQCMPTKTHLSINTPECEVKLTACSEENACQCMFLAGFTQ 420
 DB 361 RRLSDNSQWLOVSLVFTQIQCMPTKTHLSINTPECEVKLTACSEENACQCMFLAGFTQ 420

QY 421 TVNEARIPACSL 432
 DB 421 TVNEARIPACSL 432

Key Location/Qualifiers
 Misc-differences 14
 /note- "Wild type Ala substituted with Asn"

Search completed: December 14, 2002, 15:53:43
 Job time : 42 secs

PT additive, particularly for animal feeds
 XX
 PS Example 1: Page 7; 5pp; English.
 XX
 CC The present sequence is Escherichia coli acid phosphatase/phytase
 CC (appa) mutant (A13IN). The wild type appa is mutated to enhance
 CC its enzymatic activity. Phytases, a specific group of monoester
 CC phosphates, are required to initiate the release of phosphate from
 CC phytate, the major storage of phosphate in cereal foods or feeds. The
 CC mutant acid phosphatase/phytase is added to animal feed as a food
 CC additive to improve uptake of phosphate and zinc from the diet. The
 CC phytase can also be added to human food. The mutant phytase is more heat
 CC stable than phytase isolated from Aspergillus niger and safer for use
 CC in human food manufacture.
 CC Notes: the present sequence is not shown in the specification but is
 CC derived from wild appa referred as SEQ ID No. 2 (AA02644) and shown on
 CC page 6-7 of the specification.
 XX
 SS Sequence 432 AA;

Query Match 95.5% Score 2177; DP 22; Length 432;
 Best Local Similarity 97.7%; Pred. No. 6,26 211;
 Matches 422; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 MKATLIPFSLIIPITPQSAFASQSFPIKLESVVIVSRHCVAPTRATQIMQVITPAMP 60
 DB 1 MKATLIPFSLIIPITPQSAFASQSFPIKLESVVIVSRHCVAPTRATQIMQVITPAMP 60

QY 61 TWVVKIGFTTPRGFTIAYIGHYWQRIYVAQHIKPGGTCQSCQVAIIADVDPSTKTF 120
 DB 61 TWVVKIGFTTPRGFTIAYIGHYWQRIYVAQHIKPGGTCQSCQVAIIADVDPSTKTF 120

QY 121 AFAGLAPICATIVHTGATSSPHFENLKTGWAQDENANVTGAILERAGSLADEF 180
 DB 121 AFAGLAPICATIVHTGATSSPHFENLKTGWAQDENANVTGAILERAGSLADEF 180

QY 181 YQAFRELEVLPQSNLCIEPEKQESQSLTQALPSELKVSADCVSLGAVSLASMLT 240
 DB 181 YQAFRELEVLPQSNLCIEPEKQESQSLTQALPSELKVSADCVSLGAVSLASMLT 240

QY 241 ETELQQAQCMPEGWGRITDSHOWNTLSIHNAQPHLQRTPEVAKSRATPLILIKTA 300
 DB 241 ETELQQAQCMPEGWGRITDSHOWNTLSIHNAQPHLQRTPEVAKSRATPLILIKTA 300

QY 301 LIDPHPPKQAYGVITPISVLEIACHDINLANIGALELNWITLQAPENITFVEIPEEW 360
 DB 301 LIDPHPPKQAYGVITPISVLEIACHDINLANIGALELNWITLQAPENITFVEIPEEW 360

QY 361 RRLSDNSQWLOVSLVFTQIQCMPTKTHLSINTPECEVKLTACSEENACQCMFLAGFTQ 420
 DB 361 RRLSDNSQWLOVSLVFTQIQCMPTKTHLSINTPECEVKLTACSEENACQCMFLAGFTQ 420

QY 421 TVNEARIPACSL 432
 DB 421 TVNEARIPACSL 432

RESULT 15
 AA02644
 ID AA02644 standard; Protein: 432 AA.
 AC
 XX AA02644;
 DT 06-AUG-2001 (first entry)
 DE
 XX E. coli acid phosphatase/phytase (appa) mutant (A13IN).
 KW Acid phosphatase-phytase; appa; enzymatic activity; mutant; mutagen;
 KW animal food; monoester phosphate; cereal food; human food; food additive
 XX Escherichia coli.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Misc-differences 14
 FT /note- "Wild type Ala substituted with Asn"
 XX
 XX W0200146607 A1.
 XX
 XX 25 MAY 2001.
 XX
 XX 17 NOV-2000; 2000W00501622.
 XX
 XX 18 NOV 1999; 990S-0166179.
 XX (CORR) CORNELL RES FOUND INC.
 XX
 XX Lei X;
 XX
 XX W01; 2001 467672/38.
 XX
 XX Mutated acid phosphatase/phytase from Escherichia coli has improved
 PT enzymatic activity compared to the wild type and is useful as a food

GenCore version 5.1.3
Copyright (c) 1993 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 15:49:55 : Search time 13 seconds
(without alignments)
1391.052 Million cell updates/sec

Title: US-09-866-379a-10

Perfect score: 2279

Sequence: 1 MKAILPFLSLPLPLPSA : GFTQIVNEAFIPACSLPSHL 436

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476428 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 2182 | 95.7 | 422 | PPA_ECOLI | P05445 escherichia |
| 2 | 554 | 24.3 | 413 | AGF_ECOLI | P19926 escherichia |
| 3 | 552 | 24.2 | 413 | AGP_SALTY | Q33921 salmonella |
| 4 | 538 | 23.6 | 417 | AGP_PROPE | Q52309 providencia |
| 5 | 134.5 | 5.9 | 423 | PPAL_HUMAN | P11117 homo sapien |
| 6 | 119 | 5.2 | 423 | PPAL_FAT | P20611 rattus norv |
| 7 | 111.5 | 4.9 | 421 | PPAL_MOUSE | P24638 mus musculu |
| 8 | 111.5 | 4.9 | 457 | PAX8_MOUSE | Q00288 mus musculu |
| 9 | 109.5 | 4.8 | 479 | PHYB_ASFAW | P34755 aspergillus |
| 10 | 108.5 | 4.8 | 479 | PHYB_ASFNE | P34754 aspergillus |
| 11 | 108 | 4.7 | 386 | PPAP_HUMAN | P15309 homo sapien |
| 12 | 107.5 | 4.7 | 459 | PAX8_CANFA | P47240 canis famil |
| 13 | 105 | 4.6 | 450 | PAX8_HUMAN | Q06710 homo sapien |
| 14 | 103.5 | 4.5 | 1048 | SBCC_ECOLI | P13458 escherichia |
| 15 | 103 | 4.5 | 765 | PAVL_CACGE | Q74119 arabidopsi |
| 16 | 101.5 | 4.5 | 635 | TPOR_HUMAN | P40238 homo sapien |
| 17 | 103 | 4.4 | 471 | PPAP_FAT | Q44952 escherichia |
| 18 | 98.5 | 4.3 | 381 | PPAP_RAT | P20645 rattus norv |
| 19 | 96.5 | 4.2 | 328 | PXAL_DSLAL | Q41504 pseudomonas |
| 20 | 95 | 4.2 | 2554 | TLEL_PSEMP | P13368 desophilla |
| 21 | 94 | 4.1 | 341 | TIN2_MOUSE | Q04999 mus musculu |
| 22 | 94 | 4.1 | 612 | MYLX_XNFA | Q01573 goropus lab |
| 23 | 94 | 4.1 | 755 | PPAX_CADEL | Q07549 caenorhabdi |
| 24 | 94 | 4.1 | 895 | DAG1_BOVIN | Q18738 bos taurus |
| 25 | 93.5 | 4.1 | 972 | ORC4_SCHPO | Q97794 schizosacch |
| 26 | 93 | 4.1 | 625 | TPOR_MOUSE | Q04851 mus musculu |
| 27 | 92.5 | 4.1 | 763 | PPAR_HUMAN | Q04638 homo sapien |
| 28 | 92.5 | 4.1 | 1085 | CASR_BOVIN | P35384 bos taurus |
| 29 | 91.5 | 4.0 | 291 | CEC_PFEYO | P16101 cyathobius |
| 30 | 91.5 | 4.0 | 872 | PPAL_FAT | P42335 rattus norv |
| 31 | 91 | 4.0 | 458 | PAX8_FAT | P51974 rattus norv |
| 32 | 91 | 4.0 | 551 | CAP_SCHPO | P36521 schizosacch |
| 33 | 91 | 4.0 | 638 | SKPR_HUMAN | P08240 homo sapien |

RESULT 1

| ID | PPA_ECOLI | STANDARD: | PRT: | 432 AA. |
|----|--|-----------|------|---------|
| AC | P07102 | | | |
| DT | 01-APR-1988 (rel. 07, Created) | | | |
| DI | 01-AUG-1991 (rel. 19, Last sequence update) | | | |
| DI | 16-OCT-2001 (rel. 40, Last annotation update) | | | |
| DE | Periplasmic appA protein precursor [includes: Phosphoanhydride | | | |
| DE | phosphohydrolase] (EC 3.1.3.2) (PH 2.5 acid phosphatase) (AP); 6- | | | |
| DE | phytase (EC 3.1.3.26)] | | | |
| GN | APP A OR P0800.. | | | |
| OS | Escherichia coli. | | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; | | | |
| OC | Escherichia | | | |
| OX | NCBI_TaxID=562; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM H.A., AND SEQUENCE OF 23-29. | | | |
| RC | STRAIN-K12; | | | |
| EX | MEDLINE=90368416; PubMed=2168385; | | | |
| RA | Dassa J., Marek C., Boquet P.L.; | | | |
| RT | "The complete nucleotide sequence of the Escherichia coli gene appA | | | |
| RT | reveals significant homology between pH 2.5 acid phosphatase and | | | |
| RT | glucose-1-phosphatase." | | | |
| FL | J. Bacteriol. 172:5457-5469(1990). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-K12 / MG1655; | | | |
| EX | MEDLINE=97436617; PubMed=9278503; | | | |
| RA | Blattner F.R., Plunkett G. III, Beach C.A., Feller N.T., Burland V., | | | |
| RT | Riley M., Collado-Vides J., Glasner J.P., Rode C.K., Mayhew G.P., | | | |
| RA | Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., | | | |
| RT | Mau B., Shao Y.; | | | |
| RT | "The complete genome sequence of Escherichia coli K12." | | | |
| RL | Science 277:1453-1474(1997). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-K12; | | | |
| EX | MEDLINE=97662202; PubMed 8905232; | | | |
| RA | Oshima T., Alba B., Baba T., Fujita K., Hayashi K., Honjo A., | | | |
| RT | Ikemoto K., Itoh T., Katsuma N., Kitayama M., Kohno K., Makino K., | | | |
| RA | Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Miyouchi K., | | | |
| RT | Mori H., Motomizu K., Nakamura Y., Nakamoto H., Nishio T., Saito H., | | | |
| RA | Sampei G., Seki Y., Takami H., Takemoto K., Wada C., Yamamoto Y., | | | |
| RT | Yano M., Horiuchi T.; | | | |
| RT | "A 719-kb DNA sequence of the Escherichia coli K12 genome | | | |
| RT | with 6 x 10 ⁶ bp of the 12.7 Mb region containing the | | | |
| RL | DNA Res. 3:137-155(1996). | | | |
| RN | [4] | | | |
| RP | SEQUENCE OF 1-112 FROM N.A. | | | |
| EX | MEDLINE 87271766; PubMed 3038201; | | | |
| RA | Tonati E., Danchin A.; | | | |
| RT | "The structure of the promoter and amino terminal region of the pH | | | |
| RT | 2.5 acid phosphatase structural gene (appA) of E. coli: a negative | | | |
| RT | control of transcription mediated by cyclic AMP." | | | |
| RL | Biochimie 69:215-221(1987). | | | |
| RN | [5] | | | |

ALIGNMENTS

RP SEQUENCE OF 1-17 FROM N.A.
 RC STRAIN K12.
 RX MEDLINE 92049241; PubMed 1658595;
 RA Tassa J., Eshti H., Marek C., Hoon M., Kletter-Bontemps M.,
 Requet P.L.:
 RI "A new oxygen-regulated operon in *Escherichia coli* comprises the
 RI genes for a putative thiol cytochrome oxidase and for pH 2.5 acid
 RI phosphatase (appa)."
 RL Mol. Gen. Genet. 229: 441-452(1991).
 [6]
 RN CHARACTERIZATION AND SEQUENCE OF 23-45.
 RP Greiner R., Jany K.D.:
 RI "Characterization of a phytase from *Escherichia coli*."
 RL Biol. Chem. Hoppe-Seyler 372:664-665(1991).
 [7]
 RN CHARACTERIZATION AND SEQUENCE OF 23-45.
 RP MEDLINE 92256556; PubMed 849749;
 RA Greiner R., Kottetzy U., Jany K.D.:
 RI "Purification and characterization of two phytases from *Escherichia*
 RI coli."
 RL Arch. Biochem. Biophys. 303:107-113(1993).
 [8]
 RN CHARACTERIZATION.
 RP MEDLINE 20161462; PubMed 10696472;
 RA Golovan S., Wong C., Zhang J., Forsberg C.W.:
 RI "Characterization and overproduction of the *Escherichia coli* appA
 RI encoded bifunctional enzyme that exhibits both phytase and acid
 RI phosphatase activities."
 RL Can. J. Microbiol. 46:59-71(2000).
 [9]
 RN MUTAGENESIS.
 RP MEDLINE 93054596; PubMed 1429631;
 RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
 van Elteren R.L.:
 RI "Overexpression, site-directed mutagenesis, and mechanism of
 RI *Escherichia coli* acid phosphatase."
 RL J. Biol. Chem. 267:22830-22836(1992).
 [10]
 RN X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
 RP MEDLINE 20122624; PubMed 10655611;
 RA Lim D., Golovan S., Forsberg C.W., Jin Z.:
 RI "Crystal structures of *Escherichia coli* phytase and its complex with
 RI phytate."
 RL Nat. Struct. Biol. 7:108-114(2000).
 CC 1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O -> an
 CC alcohol + phosphate.
 CC 1- CATALYTIC ACTIVITY: Myo-inositol hexakisphosphate + phosphate.
 CC 1- SUBSTRATE MONOMER.
 CC 1- SUBCELLULAR LOCATION: Periplasmic.
 CC 1- INDUCTION: IN ADDITION TO CAMP-MEDIATED CONTROL, THIS ENZYME IS
 CC INDUCED WHEN BACTERIAL CULTURES REACH STATIONARY PHASE. ITS
 CC SYNTHESIS IS REPRESSION BY PHOSPHATE STARVATION OR A SHIFT FROM
 CC AEROBIC TO ANAEROBIC CONDITIONS.
 CC 1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, updated, or for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M58708; AAT2086.1;
 DR EMBL: AF000200; AAC74065.1;
 DR EMBL: D90745; BAA55745.1;
 DR EMBL: X05471; CAA23041.1;
 DR EMBL: M58708; NOT ANNOTATED CDS.
 DR EMBL: S63811; AAB20286.1;
 DR EMBL: A26534; A26534.
 DR EMBL: B46734; B46734.
 DR EMBL: S17960; S17960.

DR PIR: S18018; S18018.
 DR PIR: S33278; S33278.
 DR PDB: 1DKL; (4-AUG-00).
 DR PDB: 1DKM; (2-AUG-00).
 DR PDB: 1DKN; (4-AUG-00).
 DR PDB: 1DKO; (4-AUG-00).
 DR PDB: 1DKP; (4-AUG-00).
 DR PDB: 1DKQ; (4-AUG-00).
 DR PDB: 1DKR; (4-AUG-00).
 DR SWISS-26PAGE: P07102; (full).
 DR EcoGene: E310049; appA.
 DR InterPro: IIR000560; HisAc phosphatase.
 DR Pfam: PF004328; acid_phosphatase_1.
 DR PROSITE: PSC0616; HIS-ACID PHOSPHAT 1; 1.
 DR PROSITE: PSC0778; HIS-ACID PHOSPHAT 2; 1.
 KW Hydrolase; Periplasmic; Signal; Multimeric (alpha 2); Phosphatase
 KW Complete proteome.
 FT SIGNAL 1..22
 FT CHAIN 23..432
 FT ACT_SITE 39..49
 FT ACT_SITE 326..326
 FT DISULFID 99..130
 FT DISULFID 155..430
 FT DISULFID 200..210
 FT DISULFID 404..413
 FT CONFLICT 51..66
 FT CONFLICT 75..76
 FT CONFLICT 112..112
 FT CONFLICT D->S (IN REF. 4).
 FT SEQUENCE 432 AA; 47056 MW; 651066679177F11 Check4;
 Query Match 95.7%; Score 2182; DB 1; Length 432.
 Best local Similarity 98.1%; Pos. 170; Indels 0; Gaps 0.
 Matches 424; Conservative 0; Mismatches 8.
 QY 1 MEALLFEELSLPLTPQSAPFAGSEPELELESVVIVSRVVRAPVKALQIMQWVPLAWP 60
 DE 1 MEALLFEELSLPLTPQSAPFAGSEPELELESVVIVSRVVRAPVKALQIMQWVPLAWP 60
 QY 61 TWPVKGLTPFGELIAYLCHYWGCPVALPGLPKGPTQSSQAVATATATGEPKDE 120
 DE 61 TWPVKGLTPFGELIAYLCHYWGCPVALPGLPKGPTQSSQAVATATATGEPKDE 120
 QY 121 AFAAGLAPFCAIVITQADISSPPPTNPRTWQLINANVHATLTPAASLADLGH 180
 DE 121 AFAAGLAPFCAIVITQADISSPPPTNPRTWQLINANVHATLTPAASLADLGH 180
 QY 121 AFAAGLAPFCAIVITQADISSPPPTNPRTWQLINANVHATLTPAASLADLGH 180
 DE 121 AFAAGLAPFCAIVITQADISSPPPTNPRTWQLINANVHATLTPAASLADLGH 180
 QY 122 YQTAIALLIAYLRIEESKLTLEELPLALVLAALALPVALVAVVAVASMLL 240
 DE 122 YQTAIALLIAYLRIEESKLTLEELPLALVLAALALPVALVAVVAVASMLL 240
 QY 180 QQAERFEPVINFQSNQIKPKGQSSSEIQAQESQKVSAINVSEIDAVTASML 240
 DE 180 QQAERFEPVINFQSNQIKPKGQSSSEIQAQESQKVSAINVSEIDAVTASML 240
 QY 241 EETLLQAGCMPEPWETETTERWNTLGLHHAQFELTQFIVAVAFATLLELKA 300
 DE 241 EETLLQAGCMPEPWETETTERWNTLGLHHAQFELTQFIVAVAFATLLELKA 300
 QY 241 EETLLQAGCMPEPWETETTERWNTLGLHHAQFELTQFIVAVAFATLLELKA 300
 DE 241 EETLLQAGCMPEPWETETTERWNTLGLHHAQFELTQFIVAVAFATLLELKA 300
 QY 301 LTHHPVQKQAVGVILPTSVLTAGDTNLANLGALEENWLLQGLPNTFQSHIVFRW 360
 DE 301 LTHHPVQKQAVGVILPTSVLTAGDTNLANLGALEENWLLQGLPNTFQSHIVFRW 360
 QY 361 RRLSDSGWLOVIVTQGLQSMRKTPISNTPPEVKTLATLTPENAGVWGLAVTQ 420
 DE 361 RRLSDSGWLOVIVTQGLQSMRKTPISNTPPEVKTLATLTPENAGVWGLAVTQ 420
 QY 361 RRLSDSGWLOVIVTQGLQSMRKTPISNTPPEVKTLATLTPENAGVWGLAVTQ 420
 DE 361 RRLSDSGWLOVIVTQGLQSMRKTPISNTPPEVKTLATLTPENAGVWGLAVTQ 420
 QY 421 IVNEARIPAGSL 432
 DE 421 IVNEARIPAGSL 432
 QY 421 IVNEARIPAGSL 432
 DE 421 IVNEARIPAGSL 432
 RESULT 2
 AGP ECOLI
 ID AGP ECOLI STANDARD; PRT: 413 AA.
 AC P19926;
 DT 01-FEB-1991 (Ref. 17, created)
 DI 01-FEB-1991 (Ref. 17, last sequence updated)


```

SQ SEQUENCE 459 AA: 48839 MW: 400016448E56935 CP64:
Query Match 4.7%; Score 107.5; DB 1; Length 459;
Best Local Similarity 18.7%; Pred. No. 0.31;
Matches 81; Conservative 51; Mismatches 130; Indels 171; Gaps 16.

QY 26 PELKESVIVNSHGVRATKATQI-----MODVTPDAMPVWPVKLCGL-----TP 71
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
DB 27 DEVVGRPIVDLAIDAVPPDTISQLPVSHGCVSKILCPYETGSIPTVGGSKPKVAT 80
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
QY 72 RDELIAVLHWHWPQ-----RLVANGEL----- : : : : : : : : : : : :
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
DB 87 K---VVEKIDYKQKQNPIMFAMFIDKLLAGVCTNDVPSVSSINFIITKVKQOQPNLP 143
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
QY 96 KQK-----POS-----GQVAIADVDETPPYCEAF 122
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
DB 144 MDSVATKSLSRHTLIPSSAVTPSPQSUSIGSTYSINGLRTAGPSNDSKPKRMQSD 203
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
QY 123 AAGLAPICATVHTQADTSPPNLEPKLTCVQVQIDANVTVAILEPACGSIADPTCHVQ 192
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
DB 204 ----QDSQPISTNSQSSSSQPP---KIIPTDAESQHLEPLECPPEPQ-----HTP 247
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
QY 193 IATFFERVLNFGSNLTLEKQKQESLSLQALSEKVSALCVSL-----TGAV 223
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
DB 248 EAV----- : : : : : : : : : : : : : : : : : : : : : :
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
QY 244 SLASMLTEFLIQAGQMPPEPQWCRITDSHWNLTLSLINACFDLLOTPTEVAPSPATP- 292
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
DB 279 TLTPSNTEPL-----GRNLSHTQTVVADVHPISPAIFQETPEVSSSSSTPS 224
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
QY 293 ----LILIKIALIPHPQKQAVGVTLPTSLVF-----IAGHDINLANLGCALPLNMT 341
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
DB 325 SLSSSAFLDLQVQAGVPA-----CASVPPNAPFAASVYQFTQCALISQENVGGPT 379
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
QY 342 LRCQDNTIPDGE 354
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :
DB 380 LQYDPPHPIPSG 392
   || : : ||| ||| : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
PAX8_HUMAN
ID PAX8_HUMAN STANDARD: PRT: 450 AA.
AC Q06710; Q09155; Q16337; Q16338; Q16339;
DT 01-JUN-1994 (rel. 29, created)
DT 01-JUN-1994 (rel. 29, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Paired box protein Pax-8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN NCBI_TaxID=9606;
RC SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANT LEU 329
RP TISSUE=Kidney;
RX MEDLINE=93170167; PubMed=1337742;
RA Fidler A., Fickenscher H., Mundlos S., Winterpacht A., Zabel B.,
RT "PAX8, a human paired box gene: Isolation and expression in
RL developing thyroid, kidney and Wilms' tumors.";
RN development 116:611-623(1992).
RP SEQUENCE FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPLICING.
PC TISSUE=Kidney;
RX MEDLINE=94019277; PubMed 8413205;
RA Kozmik Z., Kurbatov P., Ivartler F., Bassilinder M.,
RT "Alternative splicing of Pax-8 gene transcripts is developmentally
RT regulated and generates isoforms with different transactivation
RT properties.";
PL Mol. Cell. Biol. 13:632-637(1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 3; 4 AND 5).
RX MEDLINE=95255300; PubMed 7737192,

```

```

PA Polster A., Wenzler P., Fickenscher H., Mundlos S., Winterpacht A.,
RA Abbott C., Plachov D.;
RT "Distinct functional properties of three human paired-box proteins,
RT PAX8, isoforms generated by alternative splicing in thyroid, kidney
RT and Wilms' tumors.";
PL Eur. J. Biochem. 228:899-911(1995).
RN [4]
RP VARIANTS TO PRO-40.
EX MEDLINE 21394129; PubMed 11502839;
RA Congdon T., Nguyen T. Q., Nequeira C.R., Barkley E.L., Medeiros-Neto G.,
RA Kopp P.;
RT "A novel mutation (Q40P) in PAX8 associated with congenital
RT hypothyroidism and thyroid hypoplasia: evidence for phenotypic
RT variability in mother and child.";
PL J. Clin. Endocrinol. Metab. 90:3562-3567(2001).
RN [5]
RP VARIANTS TO HIS-31 AND ARC-62.
RX MEDLINE=98250178; PubMed 9590296;
RA Mochizuki P.E., Lapi F., Klode H., Fritto M.L., Misero G., Chiovato L.,
RA Soudani A., Baserga M., Tassi V., Pinchera A., Fenzi G., Gruters A.,
RA Busslinger M., Di Lauro R.;
RT "PAX8 mutations associated with congenital hypothyroidism caused by
RT thyroid dysgenesis.";
PL Nat. Genet. 19:83-86(1998).
RN [6]
RP VARIANT TO TYR-57.
EX MEDLINE=21422501; PubMed 11232006;
RA Villain C., Pydlowski C., Dupres L., Heinrichs C., Abramowicz M.,
RA Malvaux P., Renneboog B., Parpa J., Costaquiol S., Vassart G.;
RT "Autosomal dominant transmission of congenital thyroid hypoplasia due
RT to loss-of-function mutation of PAX8.";
PL J. Clin. Endocrinol. Metab. 95:234-238(2001).
CC -!- FUNCTION: Transcription factor for the thyroid specific
CC expression of the genes exclusively expressed in the thyroid cell
CC type, maintaining the functional differentiation of such cells.
CC -!- SUBCELLULAR LOCATION: Nucleus.
CC -!- ALTERNATIVE PROTECTS: 5 isoforms, 1-55, 55-66, 66-80,
CC 4/Day8 and 5/Day80; are produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE EXCRETORY SYSTEM, THYROID
CC GLAND AND WILMS TUMORS.
CC -!- DEVELOPMENTAL STAGE: IN DEVELOPING EXCRETORY SYSTEM, BUBING
CC THYROID DIFFERENTIATION AND IN ADULT THYROID.
CC -!- DISEASE: Defects in PAX8 are a cause of thyroid dysgenesis (TD),
CC the most frequent cause of congenital hypothyroidism (CH). The
CC thyroid gland can be absent (athyreosis), ectopically located
CC (most often at the base of the tongue) and/or severely reduced in
CC size (hypoplasia).
CC -!- SIMILARITY: CONTAINS 1 PAIRED BOX DOMAIN.
CC
CC This SWISS PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/ebis/ebis/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL: X69699; -; NOT ANNOTATED_CDS.
CC EMBL: L19606; AAA03539.1; -;
CC EMBL: S77904; AAB34216.1; -;
CC EMBL: S77905; AAB34217.2; -;
CC EMBL: S77906; AAB34218.2; -;
CC FRL: A49126; A49126.
CC HSSP: P26367; 6PAX.
CC TRANSFAC: T01824; -;
CC TRANSFAC: T01825; -;
CC TRANSFAC: T01826; -;
CC TRANSFAC: T01827; -;
CC TRANSFAC: T02898; -;
CC GeneW: H0N00622; PAX8.
CC MIM: 167415; -;
CC MIM: 218700; -;
CC InterPro: IPR001523; Paired_box.

```


RT chromosomes (SMC) family proteins that cleaves hairpin DNA.".
 Proc. Natl. Acad. Sci. U.S.A. 95:7969-7974(1998).
 [8]
 RN CHARACTERIZATION.
 RX MEDLINE 99128292; PubMed 9927737;
 RA Connolly J.C., de Leon E.S., Leach D.B.;
 RT "DNA cleavage and degradation by the SbcCD protein complex from
 Escherichia coli.";
 RL Nucleic Acids Res. 27:1049-1046(1999).
 CC -!- FUNCTION: SbcCD cleaves DNA hairpin structures. These structures
 CC can inhibit DNA replication and are intermediates in certain DNA
 CC recombination reactions. The complex acts as a 3' to 5' double
 CC strand exonuclease that can open hairpins. It also has a 5'
 CC single-strand endonuclease activity.
 CC -!- SOURCE: heterodimer of sbcC and sbcD.
 CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. SMC SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X15981; CAA34104.1;
 DR EMBL: AF000146; AAC73500.1;
 DR EMBL: 073857; AAB18121.1;
 DR EMBL: M64787; AAA33473.1;
 DR PIR: J50350; HVECSC.
 DR PIR: S27548; S27548.
 DR EcoGene: EG10927; sbcC.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR003395; SMC_N.
 DR InterPro: IPR004582; SbcC.
 DR Pfam: PF02463; SMC_N; 1.
 DR TIGRFAMs: TIGR00618; sbc; 1.
 DR HydroLase; Nuclease; Exonuclease; DNA replication;
 KW DNA recombination. ATP-binding. Coiled-coil. Complete proteome;
 FT NP_BIND 37 44
 FT DOMAIN 191 402
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 438 502
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 528 648
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 671 695
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 770 880
 FT COILED COIL (POTENTIAL).
 SQ SEQUENCE 1048 AA. 118721 MW: 87833386.9824 CkC64;

Query Match 4.5% Score 103.5 DB 1. Length 1048,
 Best Local Similarity 24.0% Prod No. 2;

Matches 92; Conservation 12; X-matches 161; Indels 79; Gaps 20;

QY 19 SAFAQSEPEKLEKSVIVSRHGVRAH--TRATG--MDVTPDAMPFWVKLCGLTPRG 73
 DB 413 TALAQHAQEPPLPQ-HLVALHGVVPLVQVPLACVAVANVTG--EQTGN 469
 QY 74 GRLVAVLGVHGVRAVAT--HLEGG--PSSQVAVIIMVGRHKKVGEAFANGL-----AP 128
 DB 461 AAL-----NEMORY-----KKTQQLAVATVATICEQFARI-KTIEAQAQAQAQAQENP 507
 QY 129 ICAITVITCAVPSSEFTIFRFLKZGVQ--LDNANVTDAALLERAGSLAEFTGHYQT 183
 DB 508 LAUSTSHPAVEA-----VQALEPGVNSKLLALENE-----VKLLGEGCATLRQLDA 555
 QY 184 AFRELRLVNFQSNVCIHFFKQIPSCSIQALPSLKFVSADCVSLQCAVSLAEMLTEIF 243
 DB 556 ITRQVQVQVNEAGS-----LEQDE-----QALTQVQVAVFASLNIT-----LQPLEDQ 599
 QY 244 LLAQAGMPPQVMEVETDSHWNTIISLNAQFDLLQPTVEVAPSPATLLDLIKLITL- 302
 DB 600 PMLDAQDEHERQLRLISQHEHLCQIAAHNQ--ITQYQQQIPQPOQITTTLTGYALTL 657
 QY 303 PHPPQKQAYSVTLPTLSVLETAGHDTNLNLGGALE----LNWFLPQPPNTPGQF-LVF 357

LD 658 FQLEDEKSWIAIPEGEASWQWQENFLIAIENRQGLIPHLFIP-ASLELQEFETVVL 716
 QY 358 RHWRLSDNSQWIOVSIVPQTLLQ 381
 DB 717 FHWKQV--HQCTALHSQQQTLLQ 739
 RESULT 15
 DIVL_CAUCR
 ID DIVL_CAUCR STANIARD; PRF; 769 AA.
 AC Q9RQ09; Q9A2S2;
 DI 15-JUN-2002 (rel. 41, created)
 DT 15-JUN-2002 (rel. 41, last sequence update)
 DE Sensor protein divL (EC 2.7.3.-).
 GN divL, or CC3484.
 OS Caulobacter crescentus.
 OC Bacteria, Proteobacteria, alpha subdivision: Caulobacter group;
 CC Caulobacter;
 CX NCBI_TaxID=155992;
 RN [1]
 RN SEQTNST 155999 N.A. AMR A.1016888-155999 N SITE.
 RX SUPAIN=ALQ 19089 / CBI5;
 RX MEDLINE-20027501; PubMed-10557274;
 RA Wu J., Ohta N., Zhao J.L., Newton A.;
 RA "A novel bacterial tyrosine kinase essential for cell division and
 RT differentiation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:13068-13073(1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX SIKAIN-ALQ 19089 / CBI5;
 RX MEDLINE-21173698; PubMed-11259647;
 RA Nieman W.C., Feldblum T.V., Laub M.T., Paulsen I.F., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.P.K., Ohta N., Madock J.R.,
 RA Futscher I., Nelson W.C., Newton A., Stephens C., Phadke N.P., Ely R.,
 RA DeBoy P.T., Dodson P.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolony J.F., Smit J., Craven M.R., Kouri H., Shetty J., Berry K.,
 RA Interback T., Tran K., Wolf A., Vamathevan J., Emolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 Proc. Natl. Acad. Sci. U.S.A. 98:4146-4141(2001).
 CC -!- FUNCTION: Required for cell division and growth. It catalyzes the
 CC phosphorylation of CtrA and activates transcription in vitro of
 CC the cell cycle regulated flh promoter.
 CC -!- SUPPLEMENTARY LOCATION: Integral membrane protein.
 CC -!- PTM: Autophosphorylated.
 CC -!- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>
 CC or send an email to license@isb-sib.ch).
 CC
 DB EMBL: AF083422; AAF08344.2;
 DB EMBL: AF096907; AAK25446.1;
 DB TIGR: CC3484;
 DB SwissProt: Q9RQ09;
 DR InterPro: IPR004358; Bact_sens_pr_C.
 DR InterPro: IPR003661; His_KIN.
 DR InterPro: IPR004359; His_KIN_sig.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF02518; HATase_C_1.
 DR Pfam: PF09512; signal_1.
 DR PRINTS: PR00344; BCTELSENSOR.
 DR SMART: SM00387; HAIPase_C_1.
 DR SMART: SM00388; HSKA; 1.
 DR TIGRFAMs: TIGR00229; sensory_box_2.
 DR PROSITE: PS0109; HIS_KIN; 1.
 KW Sensory transduction; Transducer; Kinase; Transmembrane;
 KW Phosphorylation; Complete proteome.


```

Db 401 PPSNSQWQVSVLVEQTLQMPKPTLSNTPPEVWVLTIAAGVFNRAQMGVSLAQVET 420
QY 421 VNEARIPACLSRSH 435
| | | | | | | | | |
Db 421 VNEARIPACLSRSH 435

```

```

RESULT 2
US-09-866-379a-2
: Sequence 2: Application US/09/291.931A
: Patent No. 6183740
: GENERAL INFORMATION:
: APPLICANT: Kretz, Keith
: TITLE OF INVENTION: NOVEL PHYTASE
: FILE REFERENCE: 09010/024003
: CURRENT FILING DATE: 1999-04-13
: EARLIER FILING DATE: 1997-08-13
: EARLIER APPLICATION NUMBER: 09/259,214
: EARLIER FILING DATE: 1999-04-01
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 440
: TYPE: PRT
: ORGANISM: Escherichia coli
US-09-866-379a-2

```

```

Query Match 96.0% Score 2188 Db 4 Length 440
Best Local Similarity 97.7% Prod. No. 170 222
Matches 425 Conservative 0 Mismatches 10 Indels 0 Gaps 0

QY 1 MKAILIPFELSLIPITPSAFAUSEPELKLESVVIVSRBVKATIKATIMQVITDAMP 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MKAILIPFELSLIPITPSAFAUSEPELKLESVVIVSRBVKATIKATIMQVITDAMP 60

QY 61 TPVKLSLTPRGGLIAYLGHYWRQRLVADKLLPKGCPQSCQVATIDVDFPKTGR 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 TPVKLSLTPRGGLIAYLGHYWRQRLVADKLLPKGCPQSCQVATIDVDFPKTGR 120

QY 121 AFAAALAPKATIVHTQADTSSDPENFIRKGVQGLNANVTDAIIFPAGSLAPFTCH 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 AFAAALAPKATIVHTQADTSSDPENFIRKGVQGLNANVTDAIIFPAGSLAPFTCH 180

QY 181 YQAFRELERVLPQSNLTIRKQDQSNLTQALPSSELKVSAPVSLDAVITASMLT 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 YQAFRELERVLPQSNLTIRKQDQSNLTQALPSSELKVSAPVSLDAVITASMLT 240

QY 241 EFTLLQAGMPGEPGWGRITDSHWNLTILSHNAQFYLLQPTPEVAPRSATPLDLIKTA 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 EFTLLQAGMPGEPGWGRITDSHWNLTILSHNAQFYLLQPTPEVAPRSATPLDLIKTA 300

QY 301 LIPHPKQAGVTLPTSVLTIAHDTNLANIAGALELNWTLPGQDNTPPGGELVFERW 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 LIPHPKQAGVTLPTSVLTIAHDTNLANIAGALELNWTLPGQDNTPPGGELVFERW 360

QY 361 KRLSNSQWQVSVLVEQTLQMPKPTLSNTPPEVWVLTIAAGVFNRAQMGVSLAQVET 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 KRLSNSQWQVSVLVEQTLQMPKPTLSNTPPEVWVLTIAAGVFNRAQMGVSLAQVET 420

QY 421 VNEARIPACLSRSH 435
| | | | | | | | | |
Db 421 VNEARIPACLSRSH 435

```

```

RESULT 3
US-09-291-931-2
: Sequence 2: Application US/09/291.931A
: Patent No. 6190897
: GENERAL INFORMATION:

```

```

: APPLICANT: Kretz, Keith
: TITLE OF INVENTION: NOVEL PHYTASE
: FILE REFERENCE: 09010/024003
: CURRENT APPLICATION NUMBER: US/09/291.931A
: CURRENT FILING DATE: 1999-04-13
: EARLIER APPLICATION NUMBER: 09/010,798
: EARLIER FILING DATE: 1997-08-13
: EARLIER APPLICATION NUMBER: 09/259,214
: EARLIER FILING DATE: 1999-04-01
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 440
: TYPE: PRT
: ORGANISM: Escherichia coli
US-09-291-931-2

Query Match 96.0% Score 2188 Db 4 Length 440
Best Local Similarity 97.7% Prod. No. 170 222
Matches 425 Conservative 0 Mismatches 10 Indels 0 Gaps 0

QY 1 MKAILIPFELSLIPITPSAFAUSEPELKLESVVIVSRBVKATIKATIMQVITDAMP 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MKAILIPFELSLIPITPSAFAUSEPELKLESVVIVSRBVKATIKATIMQVITDAMP 60

QY 61 TPVKLSLTPRGGLIAYLGHYWRQRLVADKLLPKGCPQSCQVATIDVDFPKTGR 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 TPVKLSLTPRGGLIAYLGHYWRQRLVADKLLPKGCPQSCQVATIDVDFPKTGR 120

QY 121 AFAAALAPKATIVHTQADTSSDPENFIRKGVQGLNANVTDAIIFPAGSLAPFTCH 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 AFAAALAPKATIVHTQADTSSDPENFIRKGVQGLNANVTDAIIFPAGSLAPFTCH 180

QY 181 YQAFRELERVLPQSNLTIRKQDQSNLTQALPSSELKVSAPVSLDAVITASMLT 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 YQAFRELERVLPQSNLTIRKQDQSNLTQALPSSELKVSAPVSLDAVITASMLT 240

QY 241 EFTLLQAGMPGEPGWGRITDSHWNLTILSHNAQFYLLQPTPEVAPRSATPLDLIKTA 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 EFTLLQAGMPGEPGWGRITDSHWNLTILSHNAQFYLLQPTPEVAPRSATPLDLIKTA 300

QY 301 LIPHPKQAGVTLPTSVLTIAHDTNLANIAGALELNWTLPGQDNTPPGGELVFERW 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 LIPHPKQAGVTLPTSVLTIAHDTNLANIAGALELNWTLPGQDNTPPGGELVFERW 360

QY 361 KRLSNSQWQVSVLVEQTLQMPKPTLSNTPPEVWVLTIAAGVFNRAQMGVSLAQVET 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 KRLSNSQWQVSVLVEQTLQMPKPTLSNTPPEVWVLTIAAGVFNRAQMGVSLAQVET 420

QY 421 VNEARIPACLSRSH 435
| | | | | | | | | |
Db 421 VNEARIPACLSRSH 435

RESULT 4
US-09-866-379a-2
: Sequence 2: Application US/09/10798
: Patent No. 5876997
: GENERAL INFORMATION:
: APPLICANT: KRETZ
: TITLE OF INVENTION: NOVEL PHYTASE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson, P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: California
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/044,718
; FILING DATE: August 13, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HALL, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE TO PRIOR APPLICATIONS:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5039
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-044-718-2

```

```

Query Match      91.3%; Score 2080.5; DB 2; Length 423;
Best Local Similarity 91.8%; Pred. No. 3.5e-211;
Matches 408; Conservative 0; Mismatches 10; Indels 17; Gaps 1;

QY 1 MKALLPELSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQGVTPDAMP 60
DB 1 MKALLPELSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQGVTPDAMP 60

QY 61 TWPVKLGELTPRGSELIAYLGHYWPQRIVAQGLIKKQSPQSGQVAIIADVDERTPKTGE 120
DB 61 TWPVKLGELTPRGSELIAYLGHYWPQRIVAQGLIKKQSPQSGQVAIIADVDERTPKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSPPDLEPNLKTGVGHQDNANVTDAILEPAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSPPDLEPNLKTGVGHQDNANVTDAILEPAGGSIADFTGH 180

QY 181 YOTAFRELRVINEPQSNLQKREKQDECSLTQALPSELKVSADCVSLTCAVSLASMLT 240
DB 181 YOTAFRELRVINEPQSNLQKREKQDECSLTQALPSELKVSADCVSLTCAVSLASMLT 240

QY 241 EIEFLQAGMPFQWGPBRIETSHQWNTLILSNAGFYLQPTPEVAPSKAIPFLJLLIATA 300
DB 241 EIEFLQAGMPFQWGPBRIETSHQWNTLILSNAGFYLQPTPEVAPSKAIPFLJLLIATA 300

QY 301 LTPHPQKQAYGVTLPTSVLFIAQHDNLNANIGRALEINWTLPGQPUNTTPGGELVFERPW 360
DB 301 LTPHPQKQAYGVTLPTSVLFIAQHDNLNANIGRALEINWTLPGQPUNTTPGGELVFERPW 360

QY 361 RPLSNQWIGVSLVQTLQAMRKPTPLSTNTPGEPVKLALAPPEFPAQACMCSLAFITQ 420
DB 361 RPLSNQWIGVSLVQTLQAMRKPTPLSTNTPGEPVKLALAPPEFPAQACMCSLAFITQ 420

QY 421 IVNEAPTPAGSLPSH 435
DB 421 IVNEAPTPAGSLPSH 435

```

```

RESULT 5
US-09-044-718-3

```

```

; Sequence 3; Application US/09044718
; Patent No. 6391605
; GENERAL INFORMATION:
; APPLICANT: KOSTREMA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6

```

```

; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
; US-09-044-718-3

Query Match      5.5%; Score 126; DB 4; Length 449;
Best Local Similarity 20.9%; Pred. No. 0.0012;
Matches 102; Conservative 69; Mismatches 169; Indels 199; Gaps 26;

QY 7 PELSLLIPIPOSFAQSEPELKLESVVIVSRHGVRAPT-----KATQLMQGVTPDAMP 58
DB 26 PELSLEDELSSVSKLPK---DGRITLVQVLSRHGARYPTSSSKYKKLVIAQANATDF 82

QY 59 -----WPTWPVKLG--ELTPRGSELIAYLGHYWPQRIVAQGLIKKQSPQSGQVAII 108
DB 83 KGFALFKTYNYTLGADDLTPFGROLYNSGLKIFYRYKALARSVVP-----FIR 132

QY 109 ADVDERIKTIGFAAGL-----APDCAITVHIALISSDPLFNPLKTGV 154
DB 133 ASGSDFVIASGEKFTGEFQQAALADPGAINRAAPALSVLI--PESEIFN-----NILDHGV 186

QY 155 C-QLDNANVTDAILEPAGGSIADFTGHYWPQRIVAQGLIKKQSPQSGQVAII 214
DB 187 CTKFASOLGDEV-----AANFTALFADIR-----AKAEKH----- 218

QY 214 QALPSELKVSADCVSLTCAVSLASMLTETETLQAGMPFQWGPBRIETSHQWNTLILSN 271
DB 219 --LPGVILIDEDVSLMDMCSFDTV-----ARTSDASQLSPFQOLPT 258

QY 272 HN--AQDFLLQR -----TPE-VARSATFELLDLKI--ALTPHP 305
DB 259 HNEWKYNYLQSGIKYGYGACNPLQAPAGQIGFTNFIAPITPSVDQHTSTNLSVNP 318

QY 306 PQKQAYGVTLPTSVLFIA--GHDTNLANIGRALEINWTLPGQPUNTTPGGELVFERPW 363
DB 319 -----ATPPLNATMYVDVSHDMSVSIFFAL-----GIYNTEPLSPISVESAKEL 364

QY 364 SD-NSQW1---QVSLVQTLQAMRKPTPLSTNTPGEPVKLALAPPEFPAQACMCSLAFIT 419
DB 365 DGYSASVWVPPGARAYFETMQCKSEKPL--VRALINDRVVPLRGNDV--DKLGCKKINDV 422

QY 420 QIVNEAR 426
DB 423 KQLSWAR 429

RESULT 6
US-09-044-718-12
; Sequence 12; Application US/09044718
; Patent No. 6391605
; GENERAL INFORMATION:
; APPLICANT: KOSTREMA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus

```

US 09-044-718 12

Query Match 5.5%; Score 126; ID 4; Length 449;

Best local similarity 20.9%; Pred. No. 0.00013;

Matches 102; Conservative 69; Mismatches 166; Indels 150; Gaps 26;

QY 7 PFLSLTETPSAFASPEPEKLESVVIVSRHCVRAPT-----KATQMQJQVITPA 58
 DB 46 PFLSLTETPSAFASPEPEKLESVVIVSRHCVRAPT-----KATQMQJQVITPA 58
 QY 59 WPTWPKVIG--ELTPRGDELLAYIG--HYWQRQVAVGLLIPKCCQPGSGOVATI 108
 DB 93 KGEFAFKTYNYTLGALDITPFGEGQVNSGIRFYQYKALASVVP-----FIR 142
 QY 109 AWDERPKTKGFAAAGL-----APKCATIVHTGADTSSTDETFNPKFPGV 154
 DB 143 ASGSDVIVASGKPFIEGPDQAKLADPGCAINRAAFAVSVII PSESTFN NILLHGV 212
 QY 155 C-OLDNANVTDAILERAGGSTADTGHYQATFRELKRVLPQSNLTKPKKQJESGSLT 214
 DB 197 CTKEASQAGDEV-----AANETALFAPDIR -----ARAERH 244
 QY 214 QALPSEIKVSADGVSLTGAVSLASMETETFLDQAGAMPFHWGRTIUSHWNLLSL 271
 DB 229 --LPGVLTIDEDVVSLSMOMSESTIV ARTSDASQLSPFQGLFI 266
 QY 272 HN--AQDFLLQR-----TPE VAKSRATPFLDLKLT--ALTPHP 305
 DB 269 HNEWKRYNYLQSLGKRYGYGAGNPLGAGCGTINELIAGLQSPVQDHTSINSTVSNP 328
 QY 406 PQKQRCVITPTSVLFIA--GHDTNLNLANLGALLENKWTLPQGNTPPGGELVFEPWRL 363
 DB 329 --ATPELNATMYVDFSHUNSMVSTFFAL-----GIYNGTEPLESPTSVSASAKEL 374
 QY 464 SD--NSQWT QVSLVFQTLQAMRKETLSINTPGKRVATITAGCTEPNAGGMSLAGFT 419
 DB 475 DNYSSAWVPEGAFAYETPMQKSPKPEPL-VFALINRVVPLRGDEV-DKLGECKLNDEV 449
 QY 420 QIVNEAR 426
 DB 433 KGLSWAR 439

RESULT 7

US 08-868-435 34

Sequence 34, Application US/86868435

Patent No. 6291221

GENERAL INFORMATION:

APPLICANT: Van Loon, Adolphus

TITLE OF INVENTION: POLYPEPTIDES WITH PHYTASE ACTIVITY

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann-La Roche Inc.

STREET: 400 Kingsland Street

CITY: Nutley

STATE: New Jersey

COUNTRY: United States of America

ZIP: 07110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Polioase #110, Version #1.25

CURRENT APPLICATION DATA:

ATTORNEY/ATTORNEY NUMBER: 527,909,135

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/744,241

FILING DATE:

ATTORNEY/ATTORNEY INFORMATION:

NAME: Kass, Alan P

REGISTRATION NUMBER: 42142
 REFERENCE/BOOKET NUMBER: Case booked 9409
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (201) 245-4205
 TELEFAX: (201) 245-2464
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 465 amino acids
 TYPE: amino acid
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 FEATURES:
 NAME/KEY: misc_feature
 LOCATION: 104
 OTHER INFORMATION: /note="potential N glycosylation site"
 FEATURES:
 NAME/KEY: misc_feature
 LOCATION: 119
 OTHER INFORMATION: /note="potential N glycosylation site"
 FEATURES:
 NAME/KEY: misc_feature
 LOCATION: 205
 OTHER INFORMATION: /note="potential N glycosylation site"
 FEATURES:
 NAME/KEY: misc_feature
 LOCATION: 228
 OTHER INFORMATION: /note="potential N glycosylation site"
 FEATURES:
 NAME/KEY: misc_feature
 LOCATION: 347
 OTHER INFORMATION: /note="potential N glycosylation site"
 FEATURES:
 NAME/KEY: misc_feature
 LOCATION: 374
 OTHER INFORMATION: /note="potential N glycosylation site"
 US 08-868-435-34

Query Match

Best local similarity 20.9%; Score 126; ID 4; Length 465;

Matches 102; Conservative 69; Mismatches 166; Indels 150; Gaps 26;

QY 7 PFLSLTETPSAFASPEPEKLESVVIVSRHCVRAPT KATQMQJQVITPA 58
 DB 52 PFLSLTETPSAFASPEPEKLESVVIVSRHCVRAPT KATQMQJQVITPA 58
 QY 59 WPTWPKVIG--ELTPRGDELLAYIG--HYWQRQVAVGLLIPKCCQPGSGOVATI 108
 DB 109 KGEFAFKTYNYTLGALDITPFGEGQVNSGIRFYQYKALASVVP-----FIR 142
 QY 109 AWDERPKTKGFAAAGL-----APKCATIVHTGADTSSTDETFNPKFPGV 154
 DB 159 ASGSDVIVASGKPFIEGPDQAKLADPGCAINRAAFAVSVII PSESTFN NILLHGV 212
 QY 155 C-OLDNANVTDAILERAGGSTADTGHYQATFRELKRVLPQSNLTKPKKQJESGSLT 214
 DB 213 CTKEASQAGDEV-----AANETALFAPDIR -----ARAERH 244
 QY 214 QALPSEIKVSADGVSLTGAVSLASMETETFLDQAGAMPFHWGRTIUSHWNLLSL 271
 DB 245 --LPGVLTIDEDVVSLSMOMSESTIV ARTSDASQLSPFQGLFI 266
 QY 272 HN--AQDFLLQR-----TPE VAKSRATPFLDLKLT--ALTPHP 305
 DB 285 HNEWKRYNYLQSLGKRYGYGAGNPLGAGCGTINELIAGLQSPVQDHTSINSTVSNP 328
 QY 406 PQKQRCVITPTSVLFIA--GHDTNLNLANLGALLENKWTLPQGNTPPGGELVFEPWRL 363
 DB 435 --ATPELNATMYVDFSHUNSMVSTFFAL-----GIYNGTEPLESPTSVSASAKEL 374
 QY 464 SD--NSQWT QVSLVFQTLQAMRKETLSINTPGKRVATITAGCTEPNAGGMSLAGFT 419
 DB 475 DNYSSAWVPEGAFAYETPMQKSPKPEPL-VFALINRVVPLRGDEV-DKLGECKLNDEV 449

QY 420 QIVNEAR 426
: : :
DP 449 KGLSWAR 455

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match

```

Best Local Similarity 20.9%; Pred: No. 0.00013;
Matches 102; Conservative 69; Mismatches 166; Indels 150; Gaps 26;

QY 7 PFLLSLDLPQSAFAQSEPELKLESVVIVSPHCVRAPT-----KATQLMQDDVTEIA--- 58
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
Db 52 PFFSLDELSVSSKLLK---DCEITIVQVTSPECARYPTSSKKYKLVIAIQANAIDF 108
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
QY 59 --WPTWVVKLG--ELTPPGCELIYALC--HYWQRIIVANILLPKQSCPSGQVAIL 108
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
Db 109 KGFEAFELKTYNTILGADLTPFGEQIIVNSGTRFYQYKALARSVVP-----FIR 158
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
QY 109 ANVDEPTPKYGEAFAAGL-----APQCALIVHTQADISSPDPLENPKYKGV 154
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
Db 159 ASLSURVIASGEKEFIEFGQAKLADPCATNKAAPALSVII FUSEIEN ---NTLEKGV 212
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
QY 155 C-QIDNANVTDAILERAGGSIADFTGHYATFAFELEVLNFPASNICLKPKQDESCSLT 213
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
Db 213 CTRFEASQLGDEV-----AANFTALFADDIR-----ARAKKH----- 244
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
QY 214 QALPSSELKVSAGVSLGAVSLASMLTFITFLQAGMFTFGWCKITDSHQWRTLESI--- 271
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
Db 245 --LPGVTLTDDVVSMDMCSFDIV-----ARTSDAQLSPFQULFT 284
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
QY 272 HN--ADPLQLQP-----TFP-VARSPALILDI IKT--ALTHHP 305
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
Db 285 HNEWKKVNYITQSLGKYGVYCAQNLQPAQGLGFTNELLARLESIVGQHTSINSHVSNP 344
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
QY 306 FGFZANVTLTSTVLFTA--CHITRANVYALHLSWLTGTLQNTITPCELVIEKPRRL 363
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
Db 345 -----ATPLNATMYVDFSHDSMSYTFPAL-----GLYNCIIPDSFVSFAKEL 390
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
QY 361 SP NSQWNT--QVSLVFTQLQMGKTKPLSLNTPDGEVKFI IATFFPNAGMCSLAGFT 419
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
Db 391 DQYSASVVPVPRAPAYFETMCKCKEKEPL VRALLNSVVLHSDIV-DKLKRCIKLNFV 448
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
QY 420 QIVNEAR 426
   ||| ||| : : : : : ||| ||| : : : : : ||| : : : |||
Db 449 KG_LSWAR 455

RESULT 9
US-09-044-718-78
: Sequence 78, Application US/99044718
: Patent No. 6391605
: GENERAL INFORMATION:
: APPLICANT: KOSTREWA, Dirk
: APPLICANT: PASAMONTES, Louis
: APPLICANT: TOMSCHY, Andrea
: APPLICANT: van LOON, Adolphus
: APPLICANT: VOGEL, Kurt
: APPLICANT: WYSS, Markus
: TITLE OF INVENTION: MODIFIED PHYTASES
: FILE REFERENCE: Modified Phytases
: COPYRIGHT APPLICATION NUMRP9: US/99-044,718
: COPYRIGHT FILING DATE: 1998-03-19
: PRIOR APPLICATION NUMBER: EP 97810175.6
: PRIOR FILING DATE: 1997-03-25
: NUMREF OF SEQ ID NOS: 82
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 78
: LENGTH: 445
: TYPE: prt
: ORGANISM: Aspergillus fumigatus
US-09-044-718-78

```

| | | | | |
|-----------------------|--------------|-------------------|------|------------|
| Query Match | 5.58 | Score 126 | DB 4 | Length 465 |
| Best Local Similarity | 20.99 | Prod. No. 0.00013 | | |
| Matches 103 | Conservative | eq. Mismatches | 166 | Indels 150 |
| | | Gaps | 26 | |

7 PELSLLPLTPQSAFPAQSSEPELKLPSSVVSPPHIVRAPTI-----KATQIMQWTFDA---58
IIII| | : : : : : : : : : : : : : : : :
52 PEESLADPSLVSSMLPK-----DQRITPLVGVISPHZAPVTSSSKKKYKIVIALCANATDF108

```

59 QY --- WPTWPKLG- ELTPGDELLAYLG- HWYKQRLVANGLLPKGQDQSQVALL 108
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
109 DB KGFARFKLYNTLHMAIGLLEFGGGLVNSGKTFQRYKALARSVP-----FKR 158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
109 QY ADVDERTKRGAFAGL-----APDCAITVHTADTSSIPDLPFLPKYCV 154
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
159 DB ASNSDQVLASGEKLEFPQAKLADGATNPAAPALSVII PESEFEN-----NLDIKRV 212
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
159 QY QQLDNRVYDPALEFPAGNSIALPETHYGTATPELTVLPTQNLGLCKEFGDQSGTLL 213
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
213 DB QLEFASQIDDEV AANFIALPAPLEP-----ASAEKH----- 244
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
214 QY QALPSTFKVSNQVSLFCAVSLASMTFTFLQAGAMPPEPWPTTSCQANNTLLSL- 271
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
245 DB LPRGVTLDEWVSLIMDCSEDTV-----ARTSDASQISPECQLPT 284
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
272 QY HN--AUFLLQIR-----LPE-VAERSATPLDLIKT--ALITHP 305
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
285 DB HNEWKYNYLQSLGKYGYGACNPLGPAQIGFTNELIARLSPVODHTSTNLTVSNP 344
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
306 QY PQKQAYGVTLPTSVLFIA--GHDTNLANLGGALELNWTLPGQPDNTPPGGLVFFWRRL 363
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
345 DB ATPPLNATWYVDFSHDNSMVSTFEAL-----GLYNGTEPLSRITSVESAKEL 390
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
364 QY SD--NSQWLI---QVSLVFEQLQOMRKQKPLSLNTPPGVEKVLTLAGEEERNAQCMCSIACTP 419
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
391 DB DQYSASWVGVETAPAYFEETMGFSGGGLV--VGA-IMPVVPVHGGVIV--IPLEPKLNDPV 448
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
420 QY QLVNFAF 426
   | | | | |
449 DB KGLSWAR 455
   | | | | |

```

```

RESULT 10
US 09-044-714-80
? Sequence 80, Application US/09044714
? Patent No. 6391605
? GENERAL INFORMATION:
? APPLICANT: KOSTREWA, Dirk
? APPLICANT: PASAMONTES, Luis
? APPLICANT: TOMSCHY, Andrea
? APPLICANT: van LOON, Adolphus
? APPLICANT: Vogel, Kurt
? APPLICANT: WYSS, Markus
? TITLE OF INVENTION: MODIFIED PHYTASES
? FILE REFERENCE: Modified Phytases
? CURRENT APPLICATION NUMBER: US/09/044,714
? CURRENT FILING DATE: 1998-03-19
? PRIOR APPLICATION NUMBER: EP 9781075.6
? PRIOR FILING DATE: 1997-03-25
? NUMBER OF SEQ ID NOS: 82
? SOFTWARE: Patent In Ver. 2.1
? SEQ ID NO 80
? LENGTH: 465
? TYPE: PRT
? ORGANISM: Aspergillus fumigatus
US 09-044-714-80

```

```

QY 155 C QULNANVTIATLIRAGOSIALIFGHVYATFETTERVNTIUNENI KREKOCIESLII 113
      | : : : | : : : | : : : | : : : |
Db 213 CTRFEASQIGDEV ---- AANFIALFAHUR AALKH 144

QY 214 QALFSELKVSADCVSLIGAVSLASMLIFETLQACACMPFETWELIDSLWNLIEL 111
      | : : : | : : : | : : : | : : : |
Db 245 --LFAVTLTDEOVSLMDMSEDIV APTSLAQSTLPSQLEI 164

QY 272 HN -AOFDILOR -- TPE VARSKATPELDLKI ALLEHD 305
      | : : : | : : : | : : : | : : : |
Db 295 HNTWKYVNLSTLCKYGVYAGNLIAGATGHEINILLIALLISVQDHEINSLIVSNP 414

QY 306 PQKQAYCVLTTSVLELIA GHDTNLANAGALAEWELFELFQETTPFETVETWEEEL 363
      | : : : | : : : | : : : | : : : |
Db 345 -----ATPFLNATWVDFSHSNVSTFAL GLYNTEFESRISVESAKEL 500

QY 364 SD--NSQKI---QVSLVFQTLQMPKDTLSNIPCEVVKITATVTERNAJEMSLAGEL 419
      : : : | : : : | : : : | : : : |
Db 391 DGYSSASVDFGARAYFETMQCKSEREST--VPALINRVVPLDGTIV--DRLSGKLENDV 448
      : : : | : : : | : : : | : : : |

QY 420 QIVNEAR 426
      : : : | : : : | : : : | : : : |
Db 449 KGLSWAR 455

RESULT 11
US-09-044-718 79
: Sequence 79, Application 05709044718
: Patent No. 6,891,605
: GENERAL INFORMATION

```

```

? APPLICANT: PASARONTES, LUIS
? APPLICANT: TOMSCHY, Andrea
? APPLICANT: van Loon, Adolphus
? APPLICANT: WAGEL, Kurt
? APPLICANT: WYSS, Markus
? TITLE OF INVENTION: MODIFIED PHYTASPHS
? FILE REFERENCE: Modified Phytasphs
? CURRENT APPLICATION NUMBER: 05/09/044,718
? CURRENT FILING DATE: 1998-04-19
? PRIOR APPLICATION NUMBER: EP 97810179.4
? PRIOR FILING DATE: 1997-04-29
? NUMBER OF SEQ ID NOS: 82
? SOFTWARE: Patent In Vot., 2.1
? SEQ ID NO: 79
? LENGTH: 465
? TYPE: PRT
? ORGANISM: Aspergillus fumigatus
US-09-044-718-79

Query Match: 5.1%; Score 117; ID# 4; Length 465;
Best Local Similarity 20.7%; Prod. No. 0.0012;
Matches 101; Conservative 69; Mismatches 167; Indels 16; Gaps
07 7 PLSLLPLDIPGSAFACSEPELEWVAVSPHVPAP1 KALGMDLVHFA 08

```

Db 245 --LRGVILHEDVVSIMDMCSFDIV-----ARTSDASQLSPEFCQLFT 284
QY 272 HN--AGPDLQR-----TPE VAPSPAIPDLDLIKT--ALRHPH 305
Db 285 HNEWKKNYLSLGRKYGYGAGNPLGPAQIGFTNFIARLTRSPVQDHSTNSTLSVNP 344
QY 306 PQKQAYCVTLTSTVLFIA--CHYNLANLQGAIPNMTIPQDQNTDPCGPIVPEPWPPI 363
Db 345 -----ATPELNATMYVDFSHDSNMSVSIFFAL-----GLYNGTEELSPITSVESAKEL 390
QY 364 SD--NSQWI---QVSLVFETIQQMPKRTPTSTNTPRFEVKITIAKTEERNAQMGSLAGFT 419
Db 391 DQYSASVWVPPGAKAYFEIMCKSEKEPL--VPALLINPVPVPLHGCDV--EKLGRCKLNDV 448
QY 420 QVNEAP 426
Db 449 KGLSWAR 455
RESULT 12
US-09-044-718-81
: Sequence 81, Application US/09044718
: Patent No. 6391605
: GENERAL INFORMATION:
: APPLICANT: KOSTERBA, Dirk
: APPLICANT: PASAMONTA, Luis
: APPLICANT: TOMSCHY, Andrea
: APPLICANT: van LOON, Adolphus
: APPLICANT: VOGEL, Kurt
: APPLICANT: WYSS, Markus
: TITLE OF INVENTION: MODIFIED PHYTASES
: FILE REFERENCE: Modified phytases
: CURRENT APPLICATION NUMBER: US/09044,718
: CURRENT FILING DATE: 1998-03-19
: PRIOR APPLICATION NUMBER: FF 97810175 6
: PRIOR FILING DATE: 1997-03-25
: NUMBER OF SEQ ID NOS: 82
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 81
: TYPE: PRT
: ORGANISM: Aspergillus fumigatus
US-09-044-718-81
Query Match 5.18; Score 116; DB 4; Length 465;
Best Local Similarity 22.7%; Pred. No. 0.0015;
Matches 101; Conservative 50; Mismatches 167; Indels 150; Gaps 26;
QY 7 PPLSLILPLTPQSAFASPHETIKLESVIVIFVHEVPAPT-----KATQIMQDVTPDA--- 58
Db 52 PPSLEGEELSVSKLPE---IPPTIVLVLSFHAIFYETSSKSEKYLKLVTAIGANATDF 108
QY 59 -----WPTWPKLG--SLIPGSGELIAYLG--HYWPGPIVAVQVLPKGGPQSQVAIL 108
Db 109 KRLALALDYRLEGA--LEAL ELQQLVBSIFTFYGYKALAKRSVVF-- FIR 159
QY 159 ADMPEKTRKTKTGTAANGL AFDCATVHTQADTSSDPDLNPLKTVG 154
Db 159 ASGSDRVITASEKFFIEGQQAKLADPGATNKAAPISVLI--PSETFN-----NTLDHGV 212
QY 175 Q QLEMANVTRALPEAGGSAGTCHYLTAFPELEPVNPPQSNLTGKPKQRESQST 213
Db 213 CTKFASQIGDEV-----AANETALEAPDIR-----ARAKKH----- 244
QY 214 QALPSELKVSADCVSLTGAVSILASMLETEFLIQAGSGMPFGWGRITDSHOWNTLLSL- 271
Db 245 LFCVTLHEDVVSIMDMCSFDIV-----APTSASQLSPEFCQLFT 284
QY 272 HN--AGPDLQR-----TPE VAPSPAIPDLDLIKT--ALRHPH 305
Db 285 HNEWKKNYLSLGRKYGYGAGNPLGPAQIGFTNFIARLTRSPVQDHSTNSTLSVNP 344
QY 306 PQKQAYCVTLTSTVLFIA--CHYNLANLQGAIPNMTIPQDQNTDPCGPIVPEPWPPI 363

Db 345 -----ATPELNATMYVDFSHDSNMSVSIFFAL-----GLYNGTEELSPITSVESAKEL 390
QY 364 SD--NSQWI---QVSLVFETIQQMPKRTPTSTNTPRFEVKITIAKTEERNAQMGSLAGFT 419
Db 391 DQYSASVWVPPGAKAYFEIMCKSEKEPL--VPALLINPVPVPLHGCDV--EKLGRCKLNDV 448
QY 420 QVNEAP 426
Db 449 KGLSWAR 455
RESULT 13
US-07-923-724-2
: Sequence 2, Application US/07923724
: Patent No. 5780292
: GENERAL INFORMATION:
: APPLICANT: Nevalainen, Helena K.M.
: APPLICANT: Paloheimo, Marja T.
: APPLICANT: Miettinen-Oinonen, Arja S.K.
: APPLICANT: Torkkeli, Tuula K.
: APPLICANT: Cantrell, Michael
: APPLICANT: Piddington, Christopher S.
: APPLICANT: Ramesek, Joha A.
: APPLICANT: Turunen, Marja K.
: APPLICANT: Fagerstr m, Richard H.
: TITLE OF INVENTION: Production of Phytase Degrading Enzymes
: TITLE OF INVENTION: In Trichoderma
: NUMBER OF SEQUENCES: 66
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Steiner, Kessler, Goldstein & Fox
: STREET: 1100 New York Avenue, Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A
: ZIP: 20005
: COMPUTER PROGRAM: FORM-
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/923,724
: FILING DATE: 31-JUL-1992
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 97496,155
: FILING DATE: 19-MAR-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/044,077
: FILING DATE: 29-APR-1987
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: UK 8610600
: FILING DATE: 30-APR-1985
: ATTORNEY/AGENT INFORMATION:
: NAME: Cimbala, Michele A.
: PRACTICE NUMBER: 33,851
: REFERENCE/DOCKET NUMBER: 1050.0240004
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 479 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: Molecule TYPE: protein
US-07-923-724-2
Query Match 4.88; Score 109.5; DB 1; Length 479;
Best Local Similarity 19.2%; Pred. No. 0.0078;
Matches 91; Conservative 48; Mismatches 145; Indels 189; Gaps 23;

QY 29 KLESVIVSRHGVAPTKATGLMDQVTPDAMPWPKL-----GFLT----- 70
 DB 72 EVQVIMVKRCHGVSPSAG--KDELAIAKVVSNLEFKSHAVINWIIWYVYVNY 129
 QY 71 -----PRGGELIAY-----LCHYWRQRLVAIKLIPROGCHQSQVAILADWDE 114
 DB 130 YNAETISGVAGLIDAYNHCNKKYGHILWNGETVVEF-----SSGYGRVI-----E 178
 QY 114 KTKTCTCA-TAAGAGPCTATVTCAGTSSPTDPIPRKTCVQGLDRAVVTDAILLPAGG 172
 DB 179 TARKGEGFGYNTSTNAALNITSESEVMGADSL-----TPTCDTNDQITTCUNLT----- 229
 QY 173 STADFTGHVCTAFRELERVLNEFNSGLKRF-EKQCESLSLQALFSELKVSADCVSLIG 231
 DB 230 -----YQ-----LPGFKVAAAPINSQNMELT-----ASIVYNI-- 259
 QY 232 AVSLASMLTEFLIQAGQCHPEPCWCHITISQNNIL-----LSLH----- 272
 DB 260 MVMAST-----ELNAPFSNWINAFTQDEWVSFGYVEDLNYYVCAGPGUKNMAAVG 310
 QY 273 -----NAQFDLIQRTPEVARSKATILLDLIKTALTPHPQKAYGVTLPTISVLFIAHDIN 328
 DB 311 AVYANASLTLNQPKA-----GSLFFNFAHDIN 340
 QY 329 LANIGGALEL---NWTLP-----GQP---DNTPPGGELVFERW---RLSDNSQWIQ 371
 DB 341 LTPILAAAGLIPNEDLPDRVAFGNPYSIGNIYVPMGGHILTIERLSQATALSDEGTYVR 400
 QY 372 VSVVEFTIDMBQKPTLSLNTPTPEVWLTILACVPERNAQCMCSLACHTQIVNE 424
 DB 401 LVL-----NEAVLPNDTISGPNYS-CPLANYSTLNK 432

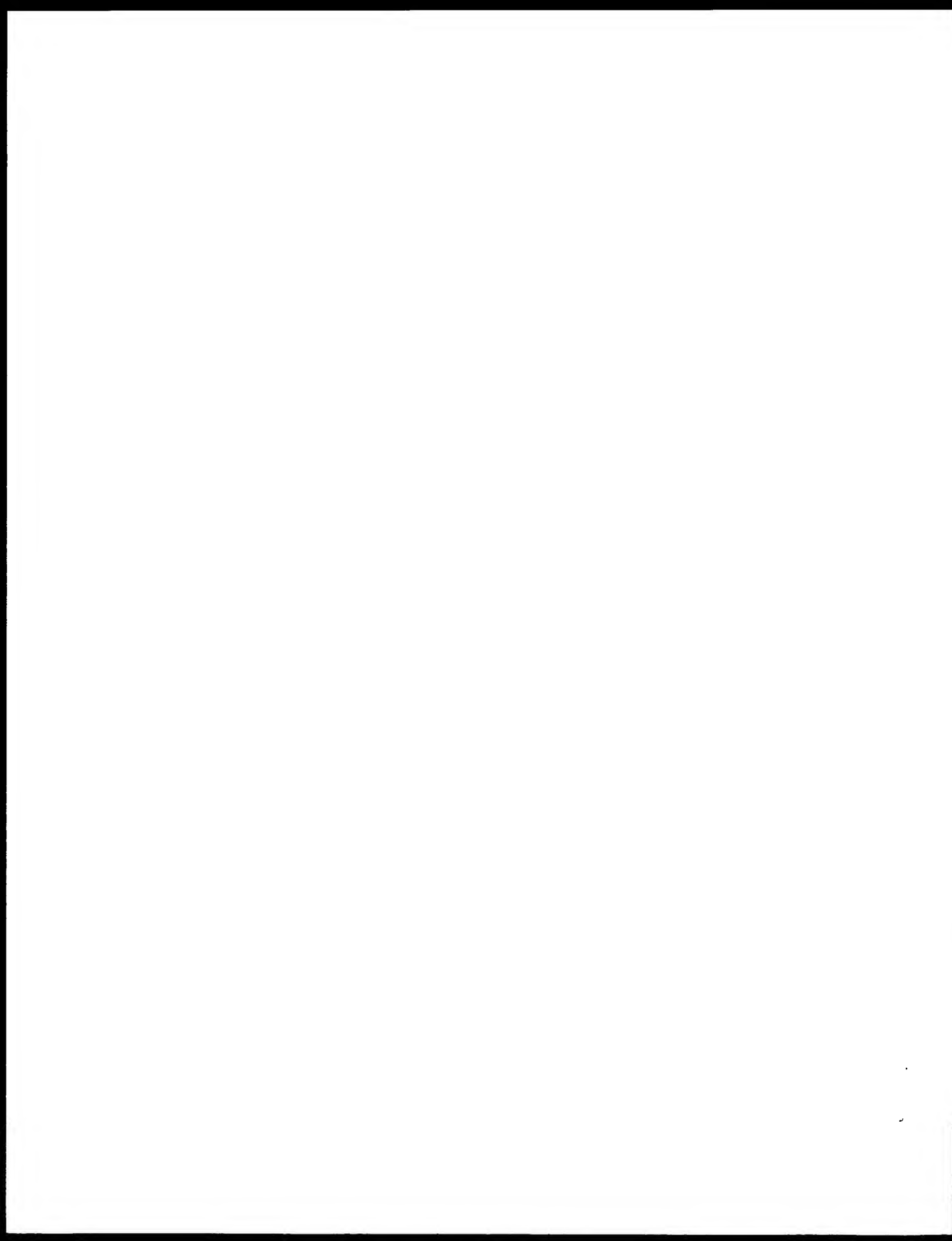
RESULT 14

US-08-609-426A-2
 Sequence 2: Application US-08-609-426A
 Patent No. 5840743
 GENERAL INFORMATION:
 APPLICANT: Nevalainen, Helena K.M.
 APPLICANT: Paloheimo, Marja T.
 APPLICANT: Miettinen-Oinonen, Arja S.K.
 APPLICANT: Torkkeli, Tuula K.
 APPLICANT: Cantrell, Michael
 APPLICANT: Piddington, Christopher S.
 APPLICANT: Ramosek, John A.
 APPLICANT: Fagerstrom, Richard B.
 APPLICANT: Houston, Christine S.
 TITLE OF INVENTION: Production of Phytase in-grading Enzymes
 NUMBER OF SEQUENCES: 69
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 STREET: 1100 New York Avenue, Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: MS-DOS/Windows
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/609,426A
 FILING DATE: 01-MAR-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/923,724
 FILING DATE: 31-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/436,155
 FILING DATE: 19-MAR-1990

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/044,077
 FILING DATE: 29-APR-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: UK 8610600
 FILING DATE: 30-APR-1986
 ATTORNEY/AGENT INFORMATION:
 NAME: Reed, Grant E.
 REGISTRATION NUMBER: P-41,264
 REFERENCE/DOCKET NUMBER: 1050.0000001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ. ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 479 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-609-426A-2
 Query Match 4.88; Score 109.5; DB:2; Locals: 479;
 Best local Similarity 19.28; Pctid: No. 0.0076;
 Matches 91; Conservative 48; Mismatches 145; Indels 189; Gaps 24;
 QY 29 KLESVIVSRHGVAPTKATGLMDQVTPDAMPWPKL-----GFLT----- 70
 DB 72 EVQVIMVKRCHGVSPSAG--KDELAIAKVVSNLEFKSHAVINWIIWYVYVNY 129
 QY 71 -----PRGGELIAY-----LCHYWRQRLVAIKLIPROGCHQSQVAILADWDE 114
 DB 130 YNAETISGVAGLIDAYNHCNKKYGHILWNGETVVEF-----SSGYGRVI-----E 178
 QY 114 KTKTCTCA-TAAGAGPCTATVTCAGTSSPTDPIPRKTCVQGLDRAVVTDAILLPAGG 172
 DB 179 TARKGEGFGYNTSTNAALNITSESEVMGADSL-----TPTCDTNDQITTCUNLT----- 229
 QY 173 STADFTGHVCTAFRELERVLNEFNSGLKRF-EKQCESLSLQALFSELKVSADCVSLIG 231
 DB 230 -----YQ-----LPGFKVAAAPINSQNMELT-----ASIVYNI-- 259
 QY 232 AVSLASMLTEFLIQAGQCHPEPCWCHITISQNNIL-----LSLH----- 272
 DB 260 MVMAST-----ELNAPFSNWINAFTQDEWVSFGYVEDLNYYVCAGPGUKNMAAVG 310
 QY 273 -----NAQFDLIQRTPEVARSKATILLDLIKTALTPHPQKAYGVTLPTISVLFIAHDIN 328
 DB 311 AVYANASLTLNQPKA-----GSLFFNFAHDIN 340
 QY 329 LANIGGALEL---NWTLP-----GQP---DNTPPGGELVFERW---RLSDNSQWIQ 371
 DB 341 LTPILAAAGLIPNEDLPDRVAFGNPYSIGNIYVPMGGHILTIERLSQATALSDEGTYVR 400
 QY 372 VSVVEFTIDMBQKPTLSLNTPTPEVWLTILACVPERNAQCMCSLACHTQIVNE 424
 DB 401 LVL-----NEAVLPNDTISGPNYS-CPLANYSTLNK 432

RESULT 15

US-08-474-652C-4
 Sequence 4: Application US/08-474-652C
 Patent No. 584286
 GENERAL INFORMATION:
 APPLICANT: Nevalainen, Helena K.M.
 APPLICANT: Paloheimo, Marja T.
 APPLICANT: Fagerstrom, Richard B.
 APPLICANT: Miettinen-Oinonen, Arja S.
 APPLICANT: Turunen, Marja K.
 APPLICANT: Ramosek, John A.
 APPLICANT: Piddington, Christopher S.
 APPLICANT: Houston, Christine S.
 APPLICANT: Cantrell, Michael A.
 TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
 US-08-474-652C-4



RA Sampaio G., Seki Y., Tsunami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horikuchi T.:
 RT "A 719-bp DNA sequence of the *Escherichia coli* K-12 *gpmA*-
 RT co-responding to the 12.7-28.0 min region on the linkage map.",
 RA DNA Res. 3:147-155(1996).
 RN [4].
 RP SEQUENCE OF 1-112 FROM N.A.
 RP MEDLINE:87271766; PubMed:3948201;
 RA Iouati E., Banchin A.:
 RT "The structure of the promoter and amino terminal region of the pH
 RT 2.5 acid phosphatase structural gene (*gppA*) of *E. coli*: a negative
 RT control of transcription mediated by cyclic AMP".
 RA Biochimie 69:215-221(1987).
 RN [5].
 RP SEQUENCE OF 1-17 FROM N.A.
 RP STRAIN:K12;
 RC MEDLINE:92049241; PubMed:1658595;
 RA Bassa J., Psihi H., Marek C., Dion M., Kieffer-Bontemps M.,
 RA Boquet P.L.:
 RT "A new oxygen-regulated operon in *Escherichia coli* comprises the
 RT phosphatase (*gppA*)."
 RA Mol. Gen. Genet. 229:341-352(1991).
 RN [6].
 RP CHARACTERIZATION, AND SEQUENCE OF 23-44.
 RP Greiner R., Jany K.D.:
 RT "Characterization of a phytase from *Escherichia coli*.";
 RA Biol. Chem. Hoppe-Seyler 372:664-665(1991).
 RN [7].
 RP CHARACTERIZATION, AND SEQUENCE OF 23-45.
 RP MEDLINE:93256556; PubMed:8487749;
 RA Greiner R., Kneidelow U., Jany K.D.:
 RT "Purification and characterization of two phytases from *Escherichia*
 RT *coli*.";
 RA Arch. Biochem. Biophys. 303:107-113(1993).
 RN [8].
 RP CHARACTERIZATION.
 RP MEDLINE:20161462; PubMed:10696472;
 RA Golovan S., Wang G., Zhang J., Fu-sheng F.W.:
 RT "Characterization and overproduction of the *Escherichia coli* *gppA*
 RT encoded bifunctional enzyme that exhibits both phytase and acid
 RT phosphatase activities".
 RA Can. J. Microbiol. 46:56-71(2000).
 RN [9].
 RP MUTAGENESIS.
 RP MEDLINE:93054596; PubMed:1429641;
 RA Costain K., Barnes E.H., Stevis P.K., Kuciel K., Zhou M.-M.,
 RA Van Etten R.L.:
 RT "Overexpression, site-directed mutagenesis, and mechanism of
 RT *Escherichia coli* acid phosphatase";
 RA J. Biol. Chem. 267:22830-22836(1992).
 RN [10].
 RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
 RP MEDLINE:20122624; PubMed:10655611;
 RA Lim B., Golovan S., Forstner F.W., Jin Z.:
 PT "Crystal structure of *Escherichia coli* phytase and its complex with
 RT phytate.";
 RA Nat. Struct. Biol. 7:108-113(2000).
 RC [11].
 RP CATALYTIC ACTIVITY: An orthophosphate monoester + H(2)O = an
 RC alcohol + phosphate.
 CC [12].
 RP CATALYTIC ACTIVITY: Myo-inositol hexakisphosphate + H(2)O = 1 myo-
 CC inositol 1,2,3,4,5-pentakisphosphate + phosphate.
 CC [13].
 RP SUBUNIT: MONOMER.
 CC [14].
 RP INDUCTION: IN ADDITION TO CAMP-MEDIATED CONTROL, THIS ENZYME IS
 CC INDUCED WHEN BACTERIAL CULTURES REACH STATIONARY PHASE; ITS
 CC SYNTHESIS IS INDUCED BY DISSOLVED PHOSPHATE STARVATION OR A SHIFT FROM
 CC AEROBIC TO ANAEROBIC CONDITIONS.
 CC [15].
 RP SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC [16].
 CC this SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its

| | | | | | |
|---|------------------------|---------------|--|------------------|-----|
| use by non-profit institutions as long as the copyright is in effect, modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.cbl-lab.org or send an email to license@csib.ch). | | | | | |
| CC | EMBL: | M5870E: | AAN72086.1; | | |
| CC | EMBL: | AF0002.00: | AAC74065.1; | | |
| CC | EMBL: | Z90745: | BAA35745.1; | | |
| CC | EMBL: | X05471: | CAG24041.1; | | |
| CC | EMBL: | M5870E: | -; NOT_ANNOTATED_CDS. | | |
| CC | EMBL: | S63811: | AAB52586.1; | | |
| CC | PIR: | A26554: | A26534. | | |
| CC | PIR: | R36733: | B36733. | | |
| CC | PIR: | S17960: | S17960. | | |
| CC | PIR: | S18018: | S18018. | | |
| CC | PIR: | S34278: | S34278. | | |
| CC | PDB: | 1DKL: | (3-AUG-00). | | |
| CC | PDB: | 1DKM: | (2-AUG-00). | | |
| CC | PDB: | 1DKN: | (3-AUG-00). | | |
| CC | PDB: | 1DKO: | (3-AUG-00). | | |
| CC | PDB: | 1DKP: | (3-AUG-00). | | |
| CC | PDB: | 1DKQ: | (3-AUG-00). | | |
| CC | PDB: | SWISS_2BBA1: | P07302; GOLF. | | |
| CC | Ecdome: | Esl0049: | appa. | | |
| CC | InterPro: | IHR00560: | HISAc-phosphatase. | | |
| CC | Tram: | P000428: | acid_phosphat_1. | | |
| CC | PROSITE: | PSC0616: | HIS_ACID_PHOSPHAT_1; 1. | | |
| CC | PROSITE: | PSC0778: | HIS_ACID_PHOSPHAT_2; 1. | | |
| KW | Hydrolase: | tertiplasmic: | Signal; Multifunctional enzyme; 4D structure; | | |
| KW | Complete proteome: | | | | |
| FT | SIGNAL | 1..22 | | | |
| FT | CHAIN | 23..432 | PERIPLASMIC AREA PROTEIN. | | |
| FT | ACT_SITE | 39..39 | NUTRIENT ACCEPTOR. | | |
| FT | ACT_SITE | 326..326 | PROTON DONOR. | | |
| FT | DISELFID | 99..130 | | | |
| FT | DISELFID | 155..430 | | | |
| FT | DISELFID | 203..210 | | | |
| FT | DISELFID | 404..413 | | | |
| FT | CONFILIP | 51..56 | MAJORITY OF THE INTRINSICALLY DISORDERED REGION (IN REF. 3). | | |
| FT | CONFLICT | 75..76 | EL. CDV (IN REF. 4) | | |
| FT | CONFLICT | 112..112 | D. S. (IN REF. 4) | | |
| FT | SEQUENCE | 432 AA: | 6510666603701136094. | | |
| SC | | | | | |
| Alignment Scores: | | | | | |
| | Pred. No.: | | 3..640..174 | Length: | 4% |
| | Score: | | 2192.00 | Matches: | 4.4 |
| | Percent Similarity: | | 98.15% | Conserved Sites: | 1 |
| | Best Local Similarity: | | 98.15% | Mismatches: | 8 |
| | Query Match: | | 94..50% | Indels: | 0 |
| | ID#: | | 1 | Gaps: | 0 |
| US-09-866-379A-9 (1..1308) X PTA ECOLI (1..432) | | | | | |

[illegible]

type, maintaining the functional differentiation of such cells.

-1- SUBCELLULAR LOCATION: Nuclear.

-1- ALTERNATIVE PRODUCTS: 5 isoforms: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000.

-1- TISSUE SPECIFICITY: EXPRESSED IN THE EXCRETORY SYSTEM, THYROID GLAND AND WILMS TUMORS.

-1- DEVELOPMENTAL STAGE: IN DEVELOPING EXCRETORY SYSTEM, DURING THYROID DIFFERENTIATION AND IN ADULT THYROID.

-1- DISEASE: Defects in PAX8 are a cause of thyroid dysgenesis (TD), the most frequent cause of congenital hypothyroidism (CH). The thyroid gland can be absent (athyreosis), ectopically located (most often at the base of the tongue) and/or severely reduced in size (hypoplasia).

-1- SIMILARITY: CONTAINS 1 PAIRED BOX DOMAIN.

This SWISS-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See <http://www.isb.sib.ch/autodoc/vy> or send an email to license@sib.ch.

EMBL: X63699; NOT_ANNOTATED_CDS.

EMBL: L19606; AAB04539.1;

EMBL: S77904; AAB34216.1;

EMBL: S77905; AAB34217.2;

EMBL: S77906; AAB34218.2;

PIR: A91263; A91263.

OSRP: P26367; PAX8.

TRANSFAC: T01824;

TRANSFAC: T01825;

TRANSFAC: T01826;

TRANSFAC: T01827;

TRANSFAC: T02898;

Genome: HGNC:8622; PAX8.

Ref: 167415;

Ref: 218700;

InterPro: IPR001524; Paired_box

PIR: P001292; PAX: 1.

PRINTS: PR00027; PAIRBOX.

SMART: SM00451; PAX: 1.

PROSITE: PS00144; PAIRED_BOX: 1.

Paired box, developmental protein; Nuclear protein; DNA-binding;

Transcription regulation; Differentiation; Alternative splicing.

Disease mutation: Polymorphism.

DOMAIN 9 133

FT VARSPIN 415 327

FT VARSPIN 400 463

FT VARSPIN 401 450

FT VARSPIN 261 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

FT VARSPIN 260 450

J. Biol. Chem. 266:2313-2319(1991).
[14] SEQUENCE FROM N.A.
MEDLINE 8923854; PubMed 271384.
Sharief F.S., Lee H., Leidenman M.M., Lundwall A., Deaven L.L., Lee C.-L., Li S.S.-L.;
"Human prostatic acid phosphatase: cDNA cloning, gene mapping and protein sequence homology with lysosomal acid phosphatase."; Biochem. Biophys. Res. Commun. 160:79-86(1989).
[15] SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
TISSUE: Prostate;
MEDLINE 98312941; PubMed 2942194;
Vihko P., Virkkunen P., Heikku P., Reiko K., Seftin T., Huhtala M.L.;
"Molecular cloning and sequence analysis of cDNA encoding human prostatic acid phosphatase."; FEBS Lett. 246:275-281(1988).
[16] SEQUENCE FROM N.A.
TISSUE: Prostate;
MEDLINE 90370491; PubMed 2395659;
Tailor P.C., Govindan M.V., Patel P.C.;
"Nucleotide sequence of human prostatic acid phosphatase determined from a full-length cDNA clone."; Nucleic Acids Res. 18:4928-4928(1990).
[17] SEQUENCE FROM N.A.
MEDLINE 95038546; PubMed 7951074;
Sharief F.S., Li S.S.-L.;
"Nucleotide sequence of human prostatic acid phosphatase ACPP gene, including seven Alu repeats."; Biochem. Mol. Biol. Int. 33:561-565(1994).
[18] X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
MEDLINE 99024966; PubMed 9804809;
Lacount M.W., Handy G., Leblond L.;
"Structural origins of L(+)-tartarate inhibition of human prostatic acid phosphatase."; J. Biol. Chem. 273:30406-30409(1998).
[19] CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an alcohol + phosphate.
[20] SUBUNIT: HOMODIMER.
[21] SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb.sib.ch/announce> or send an email to license@isb.sib.ch).

| | | | |
|-------|--------|------------|--------|
| EMBL: | M7583 | AAA60021.1 | JOINED |
| EMBL: | M7580 | AAA60021.1 | JOINED |
| EMBL: | M7581 | AAA60021.1 | JOINED |
| EMBL: | M7582 | AAA60021.1 | JOINED |
| EMBL: | M7583 | AAA60021.1 | JOINED |
| EMBL: | M7584 | AAA60021.1 | JOINED |
| EMBL: | M7585 | AAA60021.1 | JOINED |
| EMBL: | M7586 | AAA60021.1 | JOINED |
| EMBL: | M7587 | AAA60021.1 | JOINED |
| EMBL: | M7588 | AAA60021.1 | JOINED |
| EMBL: | M4840 | AAA69694.1 | - |
| EMBL: | M4902 | AAA60022.1 | - |
| EMBL: | X5174 | CAA36432.1 | - |
| EMBL: | X5175 | CAA36473.1 | - |
| EMBL: | U07097 | AAB60640.1 | JOINED |
| EMBL: | U07084 | AAB60640.1 | JOINED |
| EMBL: | U07085 | AAB60640.1 | JOINED |
| EMBL: | U07086 | AAB60640.1 | JOINED |
| EMBL: | U07088 | AAB60640.1 | JOINED |
| EMBL: | U07091 | AAB60640.1 | JOINED |
| EMBL: | U07092 | AAB60640.1 | JOINED |

[illegible]

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd

OM nucleotide - protein search, using frame_plus_n2p model

Run on: December 13, 2002, 15:55:41, Search time 43.5 seconds
(without alignments)
5781.325 Million cell updates/sec

Title: us-09-866-379a-9
Perfect score: 2309
Sequence: 1 atgaagaacatcattatcc.....gcatttgagatcatcatata 1308

Scoring table:
BL0SUM62 Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 506448

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL frame_plus_n2p_model -DEV-xlp
-Q-Zip 2 -L 1000000 -S 1000000000 -MINMATCH=0.1 -GAPCUT=0 -LOOPEXT=0
-DB-pir_73 -QFW-task=1 -SUFFIX=tbl -TPANS=human40 edi -LIST=45
-UNITS-bits -START=1 -END=1 -MAPLX-blscum62 -TPANS=human40 edi -LIST=45
-DECALCUT 200 -THP-SUFFIX=tbl -THP-MAX=100 -THP-MIN=6 -ALIGN=15 -MODE=LOCAL
-OUTFMT pto -NCRM-ext -HEADSIZ 500 -MINLEN 0 -MAXLEN=2000000000
-USERID=us09866479 -CORN=1 -72 -FUBAT 13125002 14536-21098 -NUP=6 -ICPD=3
-NO-XLPHY -NO-MAP -LAPRODUEP -NEP-SCORES=0 -WAIT -LONGTGT -DEV-TIMEEXT=128
-WARN-TIMEOUT=40 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-Ygapop=10 -Ygapext=0.5 -DELCP=6 -DELXT=7

Database : pir_73:
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1 | 2182 | 94.5 | 432 | 2 | B36733 | acid phosphatase (|
| 2 | 2144 | 92.9 | 434 | 2 | B90770 | phosphoanhydride-P |
| 3 | 2144 | 92.9 | 444 | 2 | D85633 | hypothetical prote |
| 4 | 959.5 | 41.6 | 441 | 2 | AG0201 | acid phosphatase (|
| 5 | 560 | 24.3 | 413 | 2 | F90773 | periplasmic glucos |
| 6 | 560 | 24.3 | 413 | 2 | B85636 | periplasmic glucos |
| 7 | 554 | 24.0 | 413 | 2 | JV0087 | glucose-1-phosphat |
| 8 | 547 | 23.7 | 413 | 2 | AG0642 | glucose-1-phosphat |
| 9 | 548 | 23.3 | 417 | 2 | S27627 | glucose 1 phosphat |
| 10 | 546 | 23.2 | 414 | 2 | E87315 | periplasmic phosph |
| 11 | 147 | 5.9 | 416 | 2 | T16058 | hypothetical prote |
| 12 | 144.5 | 5.8 | 423 | 1 | S06167 | acid phosphatase (|
| 13 | 119 | 5.2 | 423 | 2 | A33395 | acid phosphatase (|
| 14 | 113 | 4.9 | 3026 | 2 | T28431 | variant surface pr |

| | | | | | | |
|----|-------|-----|------|---|--------|--------------------|
| 15 | 112.5 | 4.9 | 438 | 2 | S64682 | acid phosphatase (|
| 16 | 112.5 | 4.9 | 450 | 2 | A54429 | paired box transer |
| 17 | 112 | 4.9 | 457 | 2 | A56925 | paired box transer |
| 18 | 111.5 | 4.8 | 421 | 2 | S14742 | acid phosphatase (|
| 19 | 109.5 | 4.7 | 344 | 2 | B89130 | protein F52E1.8 [l |
| 20 | 109.5 | 4.7 | 479 | 1 | JN0890 | acid phosphatase (|
| 21 | 108.5 | 4.7 | 769 | 2 | B27681 | tyrosine kinase bi |
| 22 | 108 | 4.7 | 386 | 1 | JH0610 | acid phosphatase (|
| 23 | 107.5 | 4.7 | 459 | 2 | S52250 | 3-phytase (EC 4.1. |
| 24 | 107.5 | 4.7 | 479 | 1 | JN0716 | ryanodine receptor |
| 25 | 107 | 4.6 | 5126 | 2 | S40450 | hypothetical prote |
| 26 | 106.5 | 4.6 | 397 | 2 | C81716 | kinase-related pro |
| 27 | 106 | 4.6 | 2554 | 2 | TVF671 | exonuclease (EC 4. |
| 28 | 104 | 4.5 | 1048 | 1 | BVEGSC | secreted acid phos |
| 29 | 103.5 | 4.5 | 537 | 2 | S54770 | secreted acid phos |
| 30 | 103.5 | 4.5 | 888 | 2 | T46726 | motA protein precu |
| 31 | 103.5 | 4.5 | 1662 | 2 | T18540 | hypothetical prote |
| 32 | 101.5 | 4.4 | 356 | 2 | F75594 | hypothetical prote |
| 33 | 101.5 | 4.4 | 635 | 2 | A45266 | Mbl-P protein prec |
| 34 | 101.5 | 4.4 | 789 | 2 | AF0064 | DNA directed RNA p |
| 35 | 101.5 | 4.4 | 1214 | 2 | AG2897 | conserved hypothet |
| 36 | 101.5 | 4.4 | 1387 | 2 | A97673 | probable periplasm |
| 37 | 100.5 | 4.4 | 425 | 2 | D83186 | hypothetical prote |
| 38 | 100.5 | 4.4 | 528 | 2 | I47141 | gastric mucin (clo |
| 39 | 100.5 | 4.4 | 737 | 2 | S65758 | nitrate reductase |
| 40 | 100 | 4.3 | 508 | 2 | A45477 | cAMP response elem |
| 41 | 100 | 4.3 | 715 | 2 | G75339 | probable exodeoxyr |
| 42 | 100 | 4.3 | 971 | 2 | T42883 | hypothetical prote |
| 43 | 99.5 | 4.3 | 234 | 2 | F42696 | thrombin (EC 4.4.2 |
| 44 | 99.5 | 4.3 | 1032 | 2 | T34443 | hypothetical prote |
| 45 | 99.5 | 4.3 | 2232 | 2 | T34434 | hypothetical prote |

ALIGNMENTS

RESULT 1

B36733
A:acid phosphatase (EC 3.1.3.2) precursor - Escherichia coli (strain K12)
M:Altered name: phosphoanhydride phosphohydrolase, periplasmic phytase 6
C:Species: Escherichia coli
G:Date: 19-Apr-1991, sequence_revision 15, Apr-1991, #text-change 01-Mar-2002
C:Accession: B36733, S18018, B64839, A65544, S12960, S54278
P:Passa, J.; Marek, C.; Hoquet, P.L.
J: Bacteriol. 172, 5497-5500, 1990
A:Title: The complete nucleotide sequence of the Escherichia coli gene appA reveals
A:Reference number: A36733, MIMD:96368616, PMID:2168485
A:Accession: B36733
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-432 (DAS)
A:Cross references: GR M58708, MID:9145283, FIDN:AAA72086.1; PID:d145285
R:Greiner, P.; Jany, K.D.
Biol. Chem. Hoppe-Seyler 372, 664-665, 1991
A:Title: Characterization of a phytase from Escherichia coli.
J: Bacteriol. 172, 5497-5500, 1990
A:Reference number: S18018
A:Accession: S18018
A:Molecule type: protein
A:Residues: 23-23 (GPE)
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Petna, N.T.; Burland, V.; Riley, M.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A61720, MID:97426617, PMID:9278503
A:Accession: B64839
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-432 (BLAT)
A:Cross references: GR AF000200, GR:000096; MID:92367111; PIDN:AA74065.1; PID:d171
A:Experimental source: strain K 12, substrain M61655
R:Touati, E.; Banchin, A.
Biochimie 69, 215-221, 1987
A:Title: The structure of the promoter and amino terminal region of the pH 2.5 acti
A:Reference number: A26534, MIMD:97271766; PMID:9039201

A:Accession: A26534
 A:Molecule type: DNA
 A:Residues: 150; NCACBPBRMANAS¹65¹17¹67¹74¹90¹77¹111¹8¹47025
 A:Cross-references: GR:K05471; NID:440925; PDB:CAA290211; PDB:440927
 R:Dasso, J.; Estlin, H.; Marek, C.; Bion, M.; Kieffer-Bontemps, M.; Boquet, P.L.
 Mol. Gen. Genet. 229, 341-352, 1991
 A:Title: A new oxygen-regulated operon in *Escherichia coli* comprises the genes for a putative
 A:Reference number: S17958; MUP:02049241; PMID:1658595
 A:Accession: S17960
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-17 (DA2)
 A:Cross-references: GR:K363611; NID:4238656; PDB:AA5262861; PDB:4238655
 R:Greiner, R.; Konietzky, U.; Jany, K.D.
 Arch. Biochem. Biophys. 303, 107-113, 1993
 A:Title: Purification and characterization of two phosphatases from *Escherichia coli*.
 A:Reference number: S33278; MUP:3325556; PMID:8387749
 A:Accession: S33278
 A:Molecule type: protein
 A:Residues: 24-31; A¹33-34 (GR2)
 A:Comment: In addition to CAMP-mediated control, this enzyme is induced when bacterial
 metabolic conditions.
 A:Gene: appA

A:Map position: 25 min
 C:Keywords: monomer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric monophosphate
 F:1-22/monomer; signal sequence #status predicted <SIS>
 F:24-31/27/monomer; acid phosphatase #status predicted <MAT>
 F:48/Active site; A64 #status predicted
 F:49/Active site: His (phosphohistidine intermediate) #status predicted

Alignment Scores:
 Pred. No.: 1796 173 1600th 432
 Score: 2182.00 Matches: 424
 Percent Similarity: 98.15% Conservativity: 0
 Best Local Similarity: 98.15% Mismatches: 8
 Query Match: 94.55% Indels: 0
 DB: 2 Gaps: 0

US-09 866-379A-9 (1-1308) x B36733 (1-432)

QY 1 ATCAAAAGCATCTTAATCCATTTTATCTTCGATCGGTTAAAGCCGCAATCGCA 60
 Db 1 Met Lys Ala Ileu Ileu ProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
 QY 61 TTGCTCAGAGTACGCGGAGTCAAGCTGCAAGTCTGGGATTTGTCAGTCGTCATCGT 120
 Db 21 PheAlaGlnSerGluProGlnLeuLeuGlnSerValIleValIleValSerArgHisGly 40
 QY 121 GTGCTGCTCCAAAGAGGACGACGACGATGATGATGATGATGATGATGATGATGATGAT 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 AATCTGAGGCTGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgLysGlyGlyLeuLeuAlaLysLeu 80
 QY 241 GATCAATATCTGGGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 300
 Db 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 401 CAGCTCATCTAGTGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 Db 101 GlnSerGlyGlnValAlaIleIleValAspValAspGlnAlaThrArgLysThrGlyGln 120
 QY 461 GATTTGCTGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
 Db 121 AlaPheAlaAlaLysIleAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTCTCATCTAGTGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 Db 141 SerSerProAspProLeuPheAspProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 160

QY 481 AAGCTGAGTGGAGTGGATCTTCGATGAGGAGTGGAGTGGATCTTCGATGAGTGGATCTTCGAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaArgLysLysLysLysLysLysLysLysLysLys 180
 QY 541 TATCAAAAGTGGGTTTGGGGAATTCGCAAGTGGGTTTAAATTTTGGGTTTAAATTTTGGGTT 600
 Db 181 ArgGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnIleThrLeuSerAsnLeuLys 200
 QY 601 CTTAAGATGTGAGCAACACACACGCAAGGCTGCTTATTAAGTGGTAAATTAAGTAAATTAAG 660
 Db 201 LeuLysAlaGlnLysGlnAspThrLeuThrGlnAlaLeuProSerGlnLeu 220
 QY 661 AAGCTGAGGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720
 Db 221 LysValIleSerAlaAspAspValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATCT 780
 Db 241 GluThrLeuLeuGlnGlnAlaGlnGlyMetProThrProGlyThrProGlyThrProGlyThr 260
 QY 781 GATTCATACCATCTGCAATACCT 840
 Db 261 AspSerHisGlnTrpAsnThrLeuThrLeuThrLeuThrLeuThrLeuThrLeuThrLeu 280
 QY 841 GATTCATACCATCTGCAATACCT 900
 Db 281 ArgThrIleGlnValAlaValSerArgAlaThrProLeuLeuAspLeuLeuLysThrAla 300
 QY 901 TTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 960
 Db 301 LeuThrIleGlnProGlnLysThrAlaThrGlyValThrLeuProThrSerValThr 320
 QY 961 TTAATGAGGGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
 Db 321 PheThrAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValIleLeuLeuLeuLeu 340
 QY 1021 AAGCTTCTGCTGATGAGTGGATTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProThrGlyGlnValThrLeuProThrSerValThr 360
 QY 1081 GGTGGGTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAG 1140
 Db 361 ArgArgLeuSerAspAspSerGlnTrpIleGlnValSerLeuValPheThrThrLeuGln 380
 QY 1141 CAGATGCTGATCAAAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProThrGlyGlnValLysLeuThr 400
 QY 1201 CTGCTGATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 Db 401 LeuAlaGlnLysGlnGlnArgAlaAsnAlaGlnThrLysSerLeuAlaThrPheThrThr 1280
 QY 1261 ATCTGTAATGAGGACGCTATACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1320
 Db 421 ThrValAsnGlnAlaArgLysProAlaCysSerLeu 440
 RESULT 2
 H90770
 phosphoanhydride phosphorylase [imported] - *Escherichia coli* (strain O157:H7), outst
 C:Species: *Escherichia coli*
 C:Date: 18 Jul 2001 #sequence_revision: 18 Jul 2001 #text_change: 18 Jul 2001
 C:Accession: H90770
 R:Hayashi, T.; Makino, K.; Chaiishi, M.; Kurokawa, K.; Ishida, K.; Yokoyama, K.; Hara,
 S.; Hayashi, T.; Yasunaga, T.; Kohara, S.; Shibata, T.; Hattori, M.; Shimidzu, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: H90770
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-434 (HAY)
 A:Cross-references: GR:BA000007; PDB:BA044551; PDB:150709; PDB:150714
 A:Experimental source: Strain O157:H7, substrate: PDB: 060902

C:Genetics:

A:Gene: ECs1136

Alignment Scores:
 Pred. No.: 2,82e-170 Length: 434
 Score: 2144.00 Matches: 419
 Percent Similarity: 96.54% Conservativeness: 0
 Best Local Similarity: 96.54% Mismatches: 13
 Query Match: 92.85% Indels: 2
 DB: 2 Gaps: 1

US-09-866-379A-9 (1-1308) x H90770 (1-434)

QY 1 ATGAAGGATCTTAAATCCCAATTTTATCTCTTCGATTCGGTTAAACCCGCAATCTGCA 60
 DB 1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
 QY 61 TTGCGTCAAGT-----GAGCGGAGCTGAAGCTGGAAGCTGCTGCTGCTGCTGCTGCT 114
 DB 21 PheAlaGlnSerGluProGluProGluLeuLysLeuGluSerValIleValSerArg 40
 QY 115 CATGCTATGCTGCTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 174
 DB 41 HisGlyValArgAlaLeuThrLysAlaThrGlnLeuMetGlnAspValThrProAspAla 60
 QY 175 TGCGCAACCTGGCGGCTAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 234
 DB 61 TrpProAsnTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 80
 QY 235 TATCTCGCAATLAACTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 294
 DB 81 TyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysLysGly 100
 QY 295 TATGCTATGCTGCTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 354
 DB 101 CysProGlnProCysLeuValAlaIleLeuAlaAspValAspCysLeuGlyThrArgLysThr 120
 QY 355 GCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
 DB 121 GlyLeuAlaProAlaAlaCysLeuAlaProAspGlyAlaIleThrValHisThrGluAla 140
 QY 415 GATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474
 DB 141 AspThrSerSerProAspProLeuPheAsnTrpLeuLysLysThrGlyValCysGlnLeuAsp 160
 QY 475 AAGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 534
 DB 161 AsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThr 180
 QY 535 GCGCAATTAATGAAGCGCTTTGCGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594
 DB 181 GlyHisArgPheThrAlaPheArgTyrGlnArgValLeuAsnPheProGlnSerAsn 600
 QY 595 TTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 654
 DB 201 LeuCysLeuAsnArgGlyLysGlnAspCysSerLeuThrGlyAlaLeuProSer 220
 QY 655 GAACCTAAGCTGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714
 DB 221 GluLeuGlyValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMet 240
 QY 715 GCGAGCGACATATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774
 DB 241 LeuThrGlyIlePheLeuLeuValAlaIleLeuAlaValSerLeuProGlyTrpGlyArg 260
 QY 775 ATCAGCGATTTACACCATGGAACACCTTGGTAACTTTCGCTAATGCGGCAATTTGCTTGG 834
 DB 261 ThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeu 280
 QY 835 CTATCAAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
 DB 281 LeuGlnArgThrProGluValAlaAlaGlySerAlaIleThrProLeuLeuAspLeuIleMet 300

QY 895 ACAGCGTTACGCCCATCCACGCCCAAAACAGCGGTATGGTGTGACATTAACGATATCA 954
 DB 301 IleAlaLeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSer 420
 QY 955 GTCATCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
 DB 321 ValLeuProGlyAlaGlyHisAspThrAsnLeuAlaLeuLeuGlyValAlaLeuLeuLeu 340
 QY 1015 AACTGAGCGCTTCCCGGCTCAACCGGATAACAGCGCGCGCGAGGTGTCGCAATGGTGTGCA 1074
 DB 341 AsnTrpThrLeuProGlyGlyLeuProAspAsnThrProProGlyGlyLeuValPheGlu 460
 QY 1075 GCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
 DB 361 ArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheSerThr 480
 QY 1135 TTACAGCAGATCGCTGTGATAAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1194
 DB 381 LeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLys 400
 QY 1195 GCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1254
 DB 401 LeuThrLeuAlaGlyCysGluArgAsnAlaGlyGlyMetCysSerLeuAlaGlyPhe 420
 QY 1255 ACGCAAACTCGTAATGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1296
 DB 421 ThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 434

RESULT 3

D85633

hypothetical protein appA [imported] - Escherichia coli (strain 0157:H7, substrain
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #ext_change 14 Sep-2001
 C:Accession: D85633
 R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.;
 Iliet, L.; Grobeck, F.J.; Davis, N.W.; Tim, A.; Dinalanta, E.; Potamouitis, K.;
 Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: AB5480; MUID:21074935; PMID:11206551

A:Accession: D85633

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-444 <STO>

A:Cross-references: C:AB005174; NID:012514245; PID:AA255529.1; CS009-0300145; UN0145

A:Experimental source: strain 0157:H7, substrain EDL933

C:Genetics:

A:Gene: appA

Alignment Scores:

Pred. No.: 2,82e-170 Length: 444
 Score: 2144.00 Matches: 419
 Percent Similarity: 96.54% Conservativeness: 0
 Best Local Similarity: 96.54% Mismatches: 13
 Query Match: 92.85% Indels: 2
 DB: 2 Gaps: 1

US-09-866-379A-9 (1-1308) x D85633 (1-444)

QY 1 ATGAAGGATCTTAAATCCCAATTTTATCTCTTCGATTCGGTTAAACCCGCAATCTGCA 60
 DB 11 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 40
 QY 61 TTGCGTCAAGT-----GAGCGGAGCTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 114
 DB 31 PheAlaGlnSerGluProGluProGluLeuLysLeuGluSerValIleValSerArg 50
 QY 115 CATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 174
 DB 51 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAla 70
 QY 175 TGCGCAACCTGGCGGCTAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 234
 DB 61 TrpProAsnTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 40

Db 229 GluGlnIleValGlySerThrValAspMetAlaHisPheAlaValAsnGlnValAsnValAsp 228
QY 678 GGGAC-----TCTCTTCATTACCGGCTGGGTAAGCGTGGCGATCAATTCGTGACGGAG 723
Db 229 GlyGlnGlyThrValGlyValThrLeuSerGlyThrValAlaValSerSerThrLeuGlyGly 248
QY 724 ATATTCTCTCTCAACAT 783
Db 249 ThrPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 268
QY 784 TCAC 843
Db 289 AlaIleAsnTrpValSerLeuLeuSerLeuHisAsnAlaGlnLeuMetAlaGlySer 288
QY 844 ACCAC 903
Db 289 ThrProTyrIleAlaGlnHisGlyThrProLeuLeuGlnGlnIleAspThrAlaLeu 308
QY 904 AGCGCCCATCCAC 954
Db 309 Thr-----LeuGlnLeuAspAlaGlnGlyLeuLysLeuProIleSerAlaGlnAsn 325
QY 955 -----CT 1011
Db 326 ArgValLeuPheLeuGlyGlyHisAspThrAsnIleAlaAsnIleAlaGlyMetLeuGly 345
QY 1012 CTCACTCGACGCT 1071
Db 346 AlaAspTrpGlnLeuProGlnLeuProAspAsnThrProGlyGlyGlyLeuValPhe 365
QY 1072 GAACT 1131
Db 366 GluLeuTrpGlnAsnProAspAsnHisGlnArgTyrValAlaValLysMetPheTyrGln 385
QY 1132 ATTTTAT 1188
Db 386 ThrMetAspGlnLeuAsnAlaGlnLysLeuAspLeuLysAsnAsnProAlaGlyIle 405
QY 1189 CTCAACT 1248
Db 406 ThrSerValAlaValAlaGlyCysGluAsnAsnGlyAspAspLysLeuGlySerLeuAsp 425
QY 1249 GCTTATACGCAATCTGGGAATGAGACACACACACACACACACACACACACACACAC 1296
Db 426 ThrPheGlnLysValAlaLysValIleGluSerAlaCysHisIle 441

RESULT 5
F90773
petioplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substrate
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #ext_change 18-Jul-2001
C:Accession: F90773
F:Hayashi, Y.; Wavino, K.; Ohnishi, M.; Kurakawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
qasawata, N.; Yasunaga, T.; Kuhara, S.; Shiba, I.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome
A:Reference number: A99629; MIM:6115623; PMID:11258766
A:Accession: F90773
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-413 <HAY>
A:Cross-references: GR:BA000007; PDB:BAK4561.1; PDB:q1360616; GSPLB:GN00154
A:Experimental source: strain O157:H7, substrate RMD 050952
C:Genetics:
A:Gene: ECs1158

Alignment Scores:
Pred. No.: 2e-38 Length: 413
Score: 560.00 Matches: 141
Percent Similarity: 50.47% Conservative: 73
Best local Similarity: 33.25% Mismatches: 174
Query Match: 24.25% Indels: 36
DB: 2 Gaps: 12

US-09-866-379a-9 (1-1308) x F90773 (1-413)
QY 37 ATTTCGTAAACCCGCCAATCTGCTTCATTCGCTGACAGTGAAGTCAGAG---CTGAAGTGTGAA 94
Db 13 ValValLeuLeuAlaSerAsnAlaGlnAlaGlnIleValProIleValLysGlnLeuIle 24
QY 94 ACTTGTGTGATTTGTCAGTCTGTAATGAGTGTGCTGCTGCTCA---ACCAGGACACAGAACATG 150
Db 33 GlnValLeuMetSerArgHisAsnLeuArgAlaProLeuAlaAsnAsnGlySerVal 52
QY 151 ATGTACCATCTCACTTCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 210
Db 53 LeuGlnSerThrProAsnLysTrpGlnIleAspValProLysGlnLeuThr 72
QY 211 CCGGCGCT 270
Db 73 ThrLysGlyGlyValLeuGluValTyrMetGlyHisTyrMetArgGluTrpLeuAlaGln 92
QY 271 GAGGGATCTGCTCTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 330
Db 93 GlnGlyMetValLysSerGlyGluCysProProProAspThrValTyrAlaTyrAlaAsn 112
QY 331 GTGACGACGCTGTAATGCTGTAATGCTGTAATGCTGTAATGCTGTAATGCTGTAATGCTG 390
Db 113 SerLeuGlnArgThrValAlaTyrAlaGlnPheIleThrGlyAlaPheProGlyCys 142
QY 391 CCAATACCTGATACCTGATACCTGATACCTGATACCTGATACCTGATACCTGATACCTGAT 450
Db 133 AspIleProValHisHisGlyGlnGlyMetGlyThrMetAspProThrPheAsnProVal 152
QY 451 AAAATCTGCTGTTTCCCACTGATACCTGATACCTGATACCTGATACCTGATACCTGATACCT 501
Db 153 IleThrAsp-----AspSerAlaAlaPheSerGlnLysAlaValAlaValAlaMet 168
QY 502 GAAAG 561
Db 169 GluLysLeuSerLysLeu---GlnLeuThrAspSerTyrGlnLeu-----LysGlnLys 183
QY 562 CTGACGCGGCTGCTTAATTTTCCGCAATCAAACTTGTGCTTAAAGCTGAGAAACAGAC 621
Db 184 LeuGlnLysLeuValAsnTyrLysAspSerProAlaCys-----LysGlnLys 199
QY 622 GAAACTCTCTCATTAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681
Db 206 GlnGlnCysSerLeuValAspCysLysAsnThrPheSerAlaLysIleGlnIleGlnIle 219
QY 682 TCATTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
Db 220 GlnValSerGlyProLeuLysValGlyAsnSerLeuValAspAlaPheThrLeuGlnTyr 239
QY 742 CCACAGCGCAATTCCTG---GACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795
Db 243 TrpGlnLysLeuPheMetLysAlaValAlaTrpGlyGlnLeuLysSerGlnLeuThr 259
QY 796 AACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855
Db 260 LysValLeuSerLysSerLysSerLysSerLysSerLysSerLysSerLysSerLysSer 279
QY 855 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 915
Db 280 AlaArgAsnValAlaLysProLeuValSerTyrIleAspLysAlaLeu----- 295
QY 916 CGGCAAAACACGGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
Db 296 -----ValThrAspArgThrSerAlaProLysIleThrValLeu 408
QY 967 GCGGACAGGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1024
Db 309 ValGlyHisAspSerAsnIleAlaSerLeuLeuThrAlaLeuAspPheLysProGln 428
QY 1024 CTTCTGAGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1084

Db 429 LeuHisAspLeuAsnGluAlaThrProIleGlyGlyLysIleValPheGlnAraGTrpArg 448
 QY 1084 GGGCTAAGGATACAGCGGAGGATTCAGGTTTCGTGCTGGTTCGTTCAGACTTACAGCAG 1143
 Db 449 AspSerLysAlaAsnAraAspLeuMetLysIleIleThrValTyrGlnSerAlaGluGln 468
 QY 1144 ATGCGGATGATAAAGCGCGCTGCTGATTAATACGCGCGCCGAGAGAGHGAAGACTGACCTG 1203
 Db 469 LeuAraAsnAlaAspAlaLeuThrGlnGlnAlaProAlaGlnAraValThrLeuGluLeu 388
 QY 1204 GCGATATGTCAGACAGCGAAATGAGCAGGATGCTGCTGGCAGCGTTTACGCAATGC 1263
 Db 489 SerGlyLysProIleAspAlaAspGlyPheCysProMetAspLysPheAspSerVal 407
 QY 1264 GTGAATGACAGCA 1275
 Db 408 LeuAsnGluAla 411

RESULT 6
 B856-06
 periplasmic glucose 1-phosphatase [imported] - Escherichia coli (strain 0157:H7, substra
 C:Species: Escherichia coli
 C:Date: 16 Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: B85636
 R:Perlmutter, N.L.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, F.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouitis, K.; Apodaca,
 Nature 409, 529-534, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.
 A:Reference number: A85480; PMID:121074935; PMID:11206551
 A:Accession: B85636
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-413 <STO>
 A:Cross-references: GB:AE005174; R10041514274; F001AA055501; CSF08_0806145; U000014
 A:Experimental source: Strain 0157:H7, Substrain EDL933
 C:Genetics:
 A:Gene: adp

Alignment Scores:
 Pred. No.: 26-38 Length: 413
 Scores: 560.00 Matches: 141
 Percent Similarity: 50.47% Conservative: 73
 Best Local Similarity: 33.25% Indels: 36
 Query Match: 24.25% Gaps: 12
 DB:

US-09 866 379a-9 (1-1308) x B85636 (1-413)

QY 47 ATTCGGTAAACCGGGAATCTGGATTCGCTCAGAGTACGGGAG---CTCAAGCTGGAA 93
 Db 13 ValValLeuLeuAlaSerAsnAlaGlnAlaGlnThrValProIleGlyTyrGlnLeuGln 42
 QY 94 AGTGTGTATTTGTCAGTGTCTATGCTGTGGTGTCTCA---ACCAAGGTCAGCGCAACTG 150
 Db 33 GlnValLeuMetSerAraGlnAsnLeuAraGlnAlaProLeuAlaAsnAsnGlySerVal 52
 QY 151 ATGCAATATTCACCGACAGATGATGCCAAGCTTCGCGCTGCTAAACCTGGGTGAGCTGACA 210
 Db 53 LeuGlnGlnSerThrProAsnLysLysProIleThrAspValProGlyGlyGlnLeuThr 72
 QY 211 CCGGAGGAGGCTGAGTAACTGCTATCTGAGCATTCGCGCTGAGCGCTGGTAGCC 270
 Db 73 ThrLysGlyGlyValLeuGluValTyrMetGlyHisTyrMetAraGlnGlnTrpLeuAlaGln 92
 QY 271 GACGATGAGGCTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 430
 Db 93 GlnGlyMetValLysSerGlyGlyCysProIleProAspThrValTyrAlaTyrAlaAsn 112
 QY 431 GTCGACAGCTACCGCTGATAAAGGCAAGCTTCGCGCGCGCTGGGCTGAGCTGCTGT 490
 Db 113 SerGlnGlnAraThrValAlaThrAlaGlnPhePheIleThrGlyAlaPhePheGlyGly 142
 QY 491 GCAATACCTATACATACCGCAGAGATACCTGAGTCCAGTCCGATCCGCTGATTTAACTCTCTA 450

Db 133 AspIleIleValHisHisGlnGlnGlyMetGlyThrMetAspProIleThrPheAspProVal 112
 QY 451 AAAACCTGGGCTGTGGCAACCTGATAAACGAGAAACGAAACGAAACGAAACGAAACGAAAC 501
 Db 153 IleThrAsp-----AspSerAlaAlaPheSerGlnGlnAlaValAlaAlaMet 468
 QY 502 GACAGAGTACGAGAGGTAAATTGTGCACTTACGAGGATTAATACAAAGGAGGAGGAGGAGGAG 561
 Db 169 GlnLysLeuLeuSerLysLeu-----GlnLeuThrAspSerLysIleLeu 484
 QY 562 CAGCAAGCGGCTGCTAAATTTTCGCAAACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621
 Db 184 LeuGlnLysIleValAsnIleValLysAspSerProAlaLys 199
 QY 622 GAAAGCTTTTCATTAACTGAAATTAACGAAATTAACGAAATTAACGAAATTAACGAAATTAAC 681
 Db 200 GlnGlyCysSerLeuValAspLysAsnThrPheSerAlaLysLysIleGlnGlnGlnGln 719
 QY 682 TCATTAAACGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
 Db 220 GlyValSerGlyProLeuGlyValGlyAsnSerLeuValAspAlaPheThrLeuGlnIleThr 759
 QY 742 GACAGGAGAAAGAG-----GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 799
 Db 240 TyrGluGlyPheProMetAspGlnValAlaThrPheGlyGlnIleLysSerAspLeuIleThr 759
 QY 796 AAACAGCTGGTAAATTTTCGCAAACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855
 Db 260 LysValLeuSerLysLeuLysAsnGlyTyrGlnAspSerLeuPheThrSerProIleVal 279
 QY 856 GCGCGCAGCGCGCGCTTATTAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 915
 Db 280 AlaArgAsnValAlaLysProLeuValSerTyrIleAspLysAlaLeu----- 295
 QY 916 CCGCAAAACAGCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
 Db 296 -----ValThrAspAraThrSerAlaProLysIleThrValLeu 408
 QY 967 GCGAGACGAGATATTAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1023
 Db 309 ValGlyHisAspSerAsnIleAlaSerLeuLeuThrAlaLeuAspPheLysProIleThr 428
 QY 1024 CTTCGCGCTCAGCGGCAATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1083
 Db 329 LeuHisAspGlnAsnGluAraGlyThrProIleGlyLysIleValPheGlnAlaThrAra 448
 QY 1084 CGCTAACGATACACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1143
 Db 349 AspSerLysAlaAsnAraAspLeuMetLysIleThrIleThrValTyrGlnSerAlaGlnGln 608
 QY 1144 ATGCGTATTAATAAGCGCGCTGTAAATTAATACGCGCGCGGAGGAGGAGGAGGAGGAGGAG 1203
 Db 369 LeuAraAsnAlaAspAlaLeuThrLeuGlnAlaAlaGlnAlaValThrLeuGlnLeu 468
 QY 1204 GCGAGATGCAACAGCGCAATATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1263
 Db 389 SerGlyLys-----ProIleAspAlaAspGlyPheCysProMetAspLysPheAspVal 407
 QY 1264 GTGAATGACAGCA 1275
 Db 408 LeuAsnGluAla 411

RESULT 7

JV0087
 glucose-1-phosphatase (US 3,134,110) precursor Escherichia coli (strain K 12)
 C:Species: Escherichia coli
 C:Date: 07 Sep-1990 #sequence_revision 07 Sep-1990 #text_change 01 Mar-2002
 C:Accession: JV0087; B64841
 R:Pradel, E.; Marek, C.; Boquet, P.L.
 J. Bacteriol. 172, 902-907, 1990
 A:Title: Nucleotide sequence and transcriptional analysis of the Escherichia coli
 A:Reference number: JV0087; PMID:9014040; PMID:215660

Alignment Scores:
 Pred. No.: 2,41e-37 Length: 413
 Score: 547.00 Matches: 138
 Percent Similarity: 49.98% Conservative: 66
 Best Local Similarity: 33.74% Mismatches: 181
 Query Match: 23.69% Indels: 24
 DB: 2 Gaps: 10

US-09-866-379A-9 (1-1308) x AG0632 (1-413)

QY 64 GCTCAGAGTACAGGAGAC--CTGAAGATGGAAGTGGTCATCTCAGTGGCTCATGGT 120
 DB 22 ALaLeuThrThrProGluGlyTyrGlnLeuGlnGlnValLeuMetMetSerArgHisasn 41
 QY 121 GTATGTCGTCA---ATCAASNTATGCAACTGATGATAGATATGTCAGCTGACAGATGG 177
 DB 42 LeuArgAlaProLeuAlaAsnAsnGlyAsnValLeuAlaGlnSerThrProAsnAlaTrp 61
 QY 178 CCAACTGCGTGGTAAACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 237
 DB 62 ProAlaTrpAspValProGlyGlyGlnLeuThrThrLysGlyGlyValLeuGluValTyr 81
 QY 248 CTGGCAATATACGGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 297
 DB 82 MetGlyHisTyrThrArgGlnTrpLeuValAlaGlnGlyLeuTyrProSerGlyGlyGly 101
 QY 298 GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 357
 DB 102 ProAlaProAspThrValTyrAlaTyrAlaAsnSerLeuGlnArgThrValAlaThrAla 121
 QY 358 GAAGCTTCG 417
 DB 122 GluPheBleThrSerAlaPheProGlyGlyAspPheProValHisGlnGlnLys 141
 QY 418 AGCTCCACCCGATCGTATTTAACTCTCTAAACAGCTGGCTTGGCACTGGATTAAC 477
 DB 142 MetGlyThrMetAspProThrPheAsnProValIleThrAspAspSerAlaAlaPheArg 161
 QY 478 GCGAAGCTGACTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 537
 DB 162 GluGlnAlaValGlnAlaPhe---MetGlnGlyAsnArgSerGlnLeu----- 175
 QY 538 CATTAACAAAGCGGTTTCGGAATCGAATCGGATGCTTAATTTCCGCAATCAACATGG 597
 DB 176 HisLeuAspGlnSerTyrGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 195
 QY 598 TGGCTTAAAGTGGAGAACAG 657
 DB 196 Cys-----LysGlnGlyHisGlnGlySerLeuTyrLeuAspAlaLysAspThrPhe 217
 QY 658 CTAAAGTGGACGGGAGCTGTGTGATTACGGGTGGGTTAAGCTGGATCAATGCTG 717
 DB 212 SerAlaAsnTyrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 231
 QY 718 ACAGAGATATTCCTTCGACACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 771
 DB 242 ValAspAlaIleThrThrGlnTyrTyrGlnGlyPheProMetAspGlnValAlaTrpGly 251
 QY 772 AGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 831
 DB 252 GlyIleHisThrAspArgGlnTrpLysValLeuSerTyrLysLysAsnGlyTyrGlnAsp 271
 QY 842 TGTATCAACGACGACGATGAGGACGACGACGACGACGACGACGACGACGACGACGACG 891
 DB 272 SerLeuPheThrSerProThrValAlaGlnAsnValAlaAlaProLeuValLysTyrIle 291
 QY 892 ANAGATGCTGACGCGGATCCAGCCCAAAAAGGCTGATGCTGTGATCAATACCACT 951
 DB 292 AspLysValLeuVal-----LysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 304
 QY 952 TCAGTGGTGTGTTATGCGGACAGACATTAATCTGGCAATCTGGCGGCGGCACTGGAG 1011
 DB 93 GlnAsnLysLeuLeuAlaAspLeuLeu-----CysProThrSerAsnGln 1011

Db 304 LysValThrValLeuValGlyHisAspSerAsnIleAlaSerLeuLeuThrAlaLeuAsp 1011
 QY 1012 CTAAAC--TGGAGCTTCGAGGATAGGAGATAAAGAGAGAGAGAGAGAGAGAGAGAG 1068
 DB 324 PheLysProTyrGlnLeuHisAspGlnTyrGlnAlaThrPheLeuGlyGlyGlnVal 1043
 QY 1069 TTTCAAAATCGAGTGGTAAAGATATAAGATATAAGATATAAGATATAAGATATAAG 1128
 DB 344 PheGlnArgThrPheHisAspGlyAsnAlaGlnAlaGlnAlaGlnAlaGlnAlaGln 1063
 QY 1129 CAGATTTTAAAGCATGAGTGAATAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1188
 DB 364 GlnSerAlaValGlnLeuArgAsnAlaGlnAlaLeuThrLeuLysSerProAlaAlaArg 1093
 QY 1189 GTGAAATGAGCCCTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1248
 DB 384 ValThrLeuGluLeuLysGlyCysProVal--AspAlaAsnGlyPheCysProLeuAsp 402
 QY 1249 GGTTTTAAAGTAAATCGTGAATGAAGCA 1275
 DB 403 LysPheAspAsnValMetAsnThrAla 411

RESULT 9
 S25627
 glucose-1-phosphatase (EC 3.1.3.10) precursor - Providencia rettgeri
 C.Studies: Providencia rettgeri
 G.Date: 25-Feb 1994 #sequence_revision 10-Nov 1995 #text_change 06-Oct 1999
 G.Accession: S25627
 R.Fiorio, M.L. Chiesurin, A. Lombardi, G. Satta, G.
 submitted to the EMBL Data Library, September 1992
 A:Reference number: S25627
 A:Accession: S25627
 A:Molecule type: DNA
 A:Residues: 1,417<R3C>
 A:CROSS-references: EMBL:X68201; NID:045771; PIR:CAA4288.1; PIR:041772
 A:Experimental source: strain PW7
 C:Genetics:
 A:Gene: asp
 C:Keywords: periplasmic space; phosphohistidine; phosphoprotein; phosphor
 F:1-29/Domain; signal sequence #status predicted -S01
 F:30-417/Protein; glucose-1-phosphatase #status predicted -MA1
 F:42/Active site: His (phosphohistidine intermediate) #status predicted
 F:314/Active site: His #status predicted

Alignment Scores:
 Pred. No.: 1,46e-36 Length: 417
 Score: 548.00 Matches: 140
 Percent Similarity: 50.24% Conservative: 94
 Best Local Similarity: 30.47% Mismatches: 175
 Query Match: 23.40% Indels: 0
 DB: 2 Gaps: 11

US-09-866-379A-9 (1-1308) x S25627 (1-417)

QY 28 TCTCTTCGATTCGCTTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 97
 DB 13 AlaLeuPheAlaProIleAlaProIleMetAlaSerThrAspAsnGlnAlaAspMetVal 94
 QY 88 CTGCAAAAGTGGGTCATTTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 147
 DB 33 LeuAspGlnValLeuValLeuSerArgHisAsnLeuAlaProIleValAsnThrLeu 147
 QY 148 CTGATGAGAGATGTCAGGCTGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 207
 DB 53 IleLeuThrGlnValThrAspLysTyrProAspThrAspThrAspLysSerTyrTyrLeu 207
 QY 208 ACAGGAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 208
 DB 73 ThrThrGlnGlyValAlaLeuValValTyrMetGlyHisTyrPheAlaLeuThrPheAsp 94
 QY 259 -----GATTTGGTGGAGAGATTTTGGTAAATGGAGATGGAGATGGAGATGGAG 412
 DB 93 GlnAsnLysLeuLeuAlaAspLeuLeu-----CysProThrSerAsnGln 412

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleotide - protein search, using frame_plus_n2p model

Run on: December 13, 2002, 15:58:16 ; Search time 89 Seconds

(without alignments)
60% 396 Million cell updates/sec

Title: US-09-866-379A-9
Perfect score: 2309
Sequence: 1 atgaagagagatcttaatgc.....gcagtttgagatctcatcta 1308

Scoring table: HUS0Me2

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Deloxt 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters 134369

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MCDL-frame_plus_n2p model -DEV=xlp
-Q=/gen2/1/usp/0/spec/genus/46379/runat_13122002_144536_21088/4pp_query.fasta_1.1475
-DB=SPREMBL_21 -QMT=fastan SUFFIX=rspt -MINMATCH=0.1 -LOCAL=0 -LOOPEXT=0
-UNUS=hits -SIAP=1 -ENB=1 -MATRIX=bitsum62 -TRANS=human40.631 -LIST=45
-LOCAL=100 -THP=SCOFF=100 -THP_MAX=100 -THP_MIN=0 -ALIGN 15 -MODE LOCAL
-OUTMT=plo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09866379 -ALIGN 1.1475 -runat_13122002_144536_21088 -NPROC 6 -TYPE 3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THRFADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop 6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_orqanelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 2176 | 94.2 | 432 | Q8K610 | Q8K610 escherichia |

| ID | Q8K610 | PRELIMINARY: | PRT: | 432 AA. |
|----|--|--------------|------|---------|
| AC | Q8K610 | | | |
| DT | 01-JUN-2002 (TrEMBLrel. 21, Created) | | | |
| DT | 01-JUN-2002 (TrEMBLrel. 21, Last sequence update) | | | |
| DT | 01-JUN-2002 (TrEMBLrel. 21, Last annotation update) | | | |
| DE | Periplasmic phosphoanhydride phosphohydrolase. | | | |
| GN | APPA. | | | |
| OS | Escherichia coli. | | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; | | | |
| OC | Escherichia | | | |
| OX | NCBI_TaxID=562; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| FX | MEDLINE=40468616; PubMed=2168385; | | | |
| RA | Dassa J., Marek C., Boquet P.-L.L. ; | | | |
| RT | "The complete nucleotide sequence of the Escherichia coli gene appA reveals significant homology between pH 2.5 acid phosphatase and glucose-1 phosphatase." | | | |
| RL | J. Bacteriol. 172:5497-5500(1990). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| FX | MEDLINE=94054596; PubMed=1429641; | | | |
| RA | Ostania K., Harms E.H., Stevis P.E., Kaciel R., Zhou M.M., | | | |

ALIGNMENTS

RESULT 1

Q8K610
ID Q8K610
AC Q8K610
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN APPA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
FX MEDLINE=40468616; PubMed=2168385;
RA Dassa J., Marek C., Boquet P.-L.L. ;
RT "The complete nucleotide sequence of the Escherichia coli gene appA reveals significant homology between pH 2.5 acid phosphatase and glucose-1 phosphatase."
RL J. Bacteriol. 172:5497-5500(1990).

US-09-866-379A-9 (1-1308) x Q8RKD8 (1-432)

```
QY 1 ATGAAGGAGATTTAAATCCCATTTTATCTCTCTGATTCGGTTAAACCGCCCAATCTCCA 60
DB 1 MetLysAlaLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
QY 61 TTGGTCTGACATGACCGGAGCTTAAGCTGGAAAGCTGGTCATTCCTACCTGCTCATGGT 120
DB 21 PheAlaGlnSerGluProGluLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGGTGTCTCAACCAAGCCGACGAAATATATATATATATATATATATATATATATATAT 180
DB 41 ValAlaAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGGCGGTAAACCTGGCTGACGTGATATATATATATATATATATATATATATATATAT 240
DB 61 ThrTrpProValTyrSerGlyTyrProThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GAGATTACTGATGATATATATATATATATATATATATATATATATATATATATATATAT 300
DB 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuAlaLysGlyCysPro 100
QY 301 CAGTCTGCTCATATGCTGATATATATATATATATATATATATATATATATATATATATATAT 360
DB 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGln 120
QY 361 GCTTCGCGGATGCTGATATATATATATATATATATATATATATATATATATATATATATAT 420
DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TGCAGTCGCGATCGGTATATATATATATATATATATATATATATATATATATATATATATAT 480
DB 141 SerSerProAspProLeuPheAsuProLeuLysThrGlyValCysGlnLeuAspAla 160
QY 481 AAGTGTACTATATATATATATATATATATATATATATATATATATATATATATATATATAT 540
DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180
QY 541 TATCAAGCGCGTTCGGGAAGCTGCAAGCGTCTTAATTCGCGCAATCAACCTTCGCG 600
DB 181 ArgGlnThrAlaPheArgGlnLeuGluArgValLeuAspPheProGlnSerAsnLeuLys 200
QY 601 CTTAAACGTGAGAACACGACGAGAAAGCTGCTTATTAATTAATTAATTAATTAATTAATTA 660
DB 201 LeuLysArgGlyLysSerAspGlnSerGlySerLeuThrGlnAlaLeuLeuSerGlnLeu 220
QY 661 AAGTGTACGGGTGATGCTGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
DB 221 LysValSerAlaAspAspValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GATATATTTCTCTCTTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 780
DB 241 GlnIlePheLeuLeuGlnGlnAlaPheGlySerLeuProGlyTyrPheArgIleThr 260
QY 781 GATTACACCAATGGAAACACCTTCTTAATTTGCAIACAGGAGCAATTCATATGTCACAA 840
DB 261 AspSerHisGlnTyrPheThrLeuLeuSerLeuHisAsnAlaGlyPheLysLeuLeuGln 280
QY 841 GCGACGGGACAGTTTCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 900
DB 281 ArgThrProGluValAlaValSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGGCCCATCCACCGCAAAAACAGCGGTATGCTGTCACATACCTACTTCACGCTG 960
DB 301 LeuThrProHisProProGlnGlyGlnAlaTyrGlyValThrLeuIleThrSerValLeu 320
QY 961 TTATGCGCGGCAATATATATATATATATATATATATATATATATATATATATATATATAT 1020
DB 321 PheIleAlaGlyLysAspThrAsnLeuAlaSerLeuGlyGlyAlaLeuGlnLeuAspTrp 340
QY 1021 AGCTTTCGGGTATATATATATATATATATATATATATATATATATATATATATATATAT 1080
```

```
DB 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheLeuArgTyr 360
QY 1081 CCTGCGGCAACGATACACAGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATG 1140
DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGCTGATAAAGCGCGCTGTCATTAAATATAGCCGCGCGAGACAGTGAATATGAC 1200
DB 381 GlnMetArgAspLysThrProLeuSerLeuAsnIleProProGlyGlyGlnValLysLeuThr 400
QY 1201 CTGCGAGGATGTGACAGACGAAATGCGACGCGCATGCTGCTTCTGCAAGCTTTTAACAA 1260
DB 401 LeuAlaGlyCysGlnGluLeuArgAspAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATGCTGATGATGACGCGCATATACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
RESULT 3
Q8PKD7
ID Q8RKD7 PRELIMINARY; PRT: 432 AA.
AC Q8RKD7,
DT 01-JUN-2002 (TREMBLrel. 21, Created)
PT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN APPA.
OS Escherichia coli.
OC Bacteria; Proteobacteria, gamma subdivision, Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93054596; PubMed 2168385;
RA Dassa J., Marek C., Boquet P., Lelièvre
SA "The appA gene, encoding a serine protease, the Escherichia coli gene appA
ET reveals a significant homology between the 2.5 acid phosphatase and
PT alkaline phosphatase."
PL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=93054596; PubMed 1429631;
RA Ostalan K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
PA Van Ertten P.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RI Escherichia coli acid phosphatase."
PL J. Biol. Chem. 267:55830-55836(1992)
DP EMBL: U04743; AAA09005.1; -.
KW Hydrolase.
SV SEQUENCE 432 AA. 46971 MW. 103664 Da. AD=1045 Q8C64;
```

Alignment Scores:

```
Pred. No.: 5,32e-178 Length: 432
Score: 2176.00 Bits: 423
Percent similarity: 97.92% Conservat: 9
Best Local Similarity: 97.92% Mismatches: 9
Query Match: 94.24% Indels: 0
DB: Caps: 0
```

US-09-866-379A-9 (1-1308) x Q8RKD7 (1-432)

```
QY 1 ATGAAGGAGATTTAAATCCCATTTTATCTCTCTGATTCGGTTAAACCGCCCAATCTCCA 60
DB 1 MetLysAlaLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
QY 61 TTGGTCTGACATGACCGGAGCTTAAGCTGGAAAGCTGGTCATTCCTACCTGCTCATGGT 120
DB 21 PheAlaGlnSerGluProGluLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGGTGTCTCAACCAAGCCGACGAAATATATATATATATATATATATATATATATATATAT 180
DB 41 ValAlaAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
```



```

OY 121 TCCACGTCGATACGCTTATTAATGCTGTAATAAACTGGCGTTGGCAATGCAATAACGG 480
DB 141 SerSerProAspPheLeuPheAsnProLeuLysThrGlyValGlySerGlnLeuAspAsnAla 160
OY 181 AGCTGACCTCAACGGATCTCTGACAGAGCGAGGAGCGTCAATGCTCACTTATCGGGCAAT 540
DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
OY 541 TATCAAAACCGCTTTCGGCAATGCAAGCGGTCTTAATTTTCGCGCAATCAAACTTGTGC 600
DB 181 ArgGlnThrAlaPheArgGlnLeuGluArgValIleuAsnPheProGlnSerAsnLeuCys 200
OY 601 GTTAAACGTCASAAACAGGACCAAGAGTGTCTATTAAGCGAGGATTAACATAGAGAACTC 660
DB 201 LeuAspAsnGlnLysGlnAspGlnSerGlySerGlnThrGlnAlaLeuProGlnGlnHis 220
OY 661 AAGCTGAGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuIle 240
OY 721 GATATATTCTGCTGCAACAGCACAGGGAATGCGGAGGCGGCGGTGGGGAAGGATGACC 780
DB 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlnPheProGlnProGlyTrpGlyArgIleThr 260
OY 781 GATTCACACGACGAGGAGACGCTTTCCTAAGTTTCATATACGAGGAGGAGGAGGAGGAG 840
DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheCysLeuLeuGln 280
OY 841 GCGACGACACACGCTGCGCGGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
DB 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
OY 901 TTGACGCGCTCATCGACGCGGCAAAACAGGCGGATAGGAGGAGGAGGAGGAGGAGGAGG 960
DB 301 LeuThrProHisSerProGlnLysGlnAlaLysValIleThrLeuProHisSerValLeu 320
OY 961 TTATCGCGCGACACCATATTAATCTGCGAAAGTCGCGGCGGCGGCGGCGGCGGCGGCGG 1020
DB 321 PheIleAlaGlyHisAspThrAsnLeuAlaSerLeuGlyValIleLeuGlnIleAsnIle 340
OY 1021 AGCTTTCGCGCTGACGCGGCGGCAAAACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
DB 341 ThrLeuProGlyGlnProAspAsnThrSerProGlyGlnIleValPheGlnArgTrp 360
OY 1081 GGTGGGCTAAGCGCAATACAGGCAAGGAGATTCAGGCTTTCGAGGAGGAGGAGGAGGAGG 1140
DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
OY 1141 CAGATGCTGATATAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 381 GlnMetArgAspGlyHisThrGlnLeuSerLeuAsnThrProProGlyGlnValLysLeu 400
OY 1201 GTGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 401 LeuAlaLysGlnGlnLysAlaAsnAlaLysGlyMetArgSerLeuAlaGlyPheThrGln 420
OY 1261 ATGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
DB 421 IleValAsnGlnAlaAlaIleProAlaCysSerLeu 432

RESULT 5
QBRK09 PRELIMINARY; PKI: 432 AA.
ID QBRK09
AC QBRK09
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE periplasmic phosphoanhydride phosphohydrolase.
GN APPA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.

```

```

OX NCBI_taxID=562:
RN [1]
SEQUENCE FROM N.A.
PX MEDLINE=90368616; PubMed=2158385;
RA Dassa J., Marck C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
reveals significant homology between pH 2.5 acid phosphatase and
glucose-1-phosphatase.";
PT
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
SEQUENCE FROM N.A.
PX MEDLINE=93054596; PubMed=1429631;
RA Ostani K., Harms E.H., Stevis P.E., Kuriel R., Zhou M.M.,
RT "Overexpression, site-directed mutagenesis, and mechanism of
Escherichia coli acid phosphatase.";
PT
RL J. Biol. Chem. 267:22830-22836(1992).
DR EMBL: 193371; AAA00003.1; ...
KW Hydrolase.
SQ SEQUENCE 432 AA: 47933 MW: 9F29H9P9C368175 CRC64:
US-09-866-379A-9 (1308) x QBRK09 (1-432)
Alignment Scores:
Pred. No.: 6,48e-178 Length: 432
Score: 2175.00 Matches: 423
Percent Similarity: 98.15% Conservative: 1
Best Local Similarity: 97.92% Mismatches: 8
Query Match: 94.20% Indels: 0
DB: 2 Gaps: 0
OY 1 ATCAAGCGATCTTAATCCCAATTTTATCTCTCTGATGCTTAAAGCGGCAATGCA 60
DB 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIlePheLeuThrProSerAla 20
OY 61 TTCCTTACACTACAGCGGAGCTCAAGCTGAAAGCTGAAAGCTGAGTGTGATGCTGCTGCTGCT 120
DB 21 PheAlaGlnSerGlnProGlnLysIleProValIleValSerArgAsnGly 40
OY 121 GTCTGCTCTCAACACACGCGGACGCTGATGATGATGATGATGATGATGATGATGATGATG 180
DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
OY 181 ACCTGCGCGGCTAAACATGAGGTGACACGCGGCGGCTGGTGGAGCTAATGCTGCTGCTGCT 240
DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProAspGlyLeuLeuIleAlaTrpLeu 80
OY 241 GCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 81 GlyHisTrpArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyLysGlyPro 100
OY 301 CAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 101 GlnSerGlyGlnValAlaIleLeuAspValAspGlnArgThrArgLysThrGlyGln 120
OY 361 GACTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 121 AlaPheAlaAlaGlyLeuAlaProAspLysAlaIleThrValHisThrGlnAlaAspThr 140
OY 421 TCCACTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 141 SerSerProAspProLeuPheAsnTrpGlnArgLysThrGlyValCysGlnLeuAspAsnAla 160
OY 481 AAGCTCATCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
OY 541 TATCAAAACCGCTTTCGGCAATGCAAGCGGTCTTAATTTTCGCGCAATCAAACTTGTGC 600
DB 181 ArgGlnThrAlaPheArgGlnLeuGluArgValIleuAsnPheProGlnSerAsnLeuCys 200
OY 601 GTTAAACGTCASAAACAGGACCAAGAGTGTCTATTAAGCGAGGATTAACATAGAGAACTC 660
DB 201 LeuAspAsnGlnLysGlnAspGlnSerGlySerGlnThrGlnAlaLeuProGlnGlnHis 220
OY 661 AAGCTGAGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuIle 240
OY 721 GATATATTCTGCTGCAACAGCACAGGGAATGCGGAGGCGGCGGTGGGGAAGGATGACC 780
DB 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlnPheProGlnProGlyTrpGlyArgIleThr 260
OY 781 GATTCACACGACGAGGAGACGCTTTCCTAAGTTTCATATACGAGGAGGAGGAGGAGGAGGAG 840
DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheCysLeuLeuGln 280
OY 841 GCGACGACACACGCTGCGCGGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
DB 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
OY 901 TTGACGCGCTCATCGACGCGGCAAAACAGGCGGATAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
DB 301 LeuThrProHisSerProGlnLysGlnAlaLysValIleThrLeuProHisSerValLeu 320
OY 961 TTATCGCGCGACACCATATTAATCTGCGAAAGTCGCGGCGGCGGCGGCGGCGGCGGCGG 1020
DB 321 PheIleAlaGlyHisAspThrAsnLeuAlaSerLeuGlyValIleLeuGlnIleAsnIle 340
OY 1021 AGCTTTCGCGCTGACGCGGCGGCAAAACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
DB 341 ThrLeuProGlyGlnProAspAsnThrSerProGlyGlnIleValPheGlnArgTrp 360
OY 1081 GGTGGGCTAAGCGCAATACAGGCAAGGAGATTCAGGCTTTCGAGGAGGAGGAGGAGGAGGAGG 1140
DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
OY 1141 CAGATGCTGATATAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 381 GlnMetArgAspGlyHisThrGlnLeuSerLeuAsnThrProProGlyGlnValLysLeu 400
OY 1201 GTGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 401 LeuAlaLysGlnGlnLysAlaAsnAlaLysGlyMetArgSerLeuAlaGlyPheThrGln 420
OY 1261 ATGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
DB 421 IleValAsnGlnAlaAlaIleProAlaCysSerLeu 432

```



```
DB 281 ArqThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCGCCACCCACCGCAAAACAGCGGTATGCTGTGTCACATTACCCACTTCAGTGTG 960
DB 401 LeuThrProHisProProGluIlySGTAlaLysGlyGValThrLeuProThrSerValLeu 320
QY 902 TTTATGCGGCGACACATATCAATATGAGTAAATGCGGCGGCGGCGGCGGCGGCGGCGG 1000
DB 421 TheHleAlaGlyAlaAspThrAspSerAlaAspGlyGlyAlaGluGluLeuAspThr 340
QY 1021 AGGCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 441 ThrLeuProGlyGlyProAspSerThrProProGlyGlyGlyGluValPheGluArgTrp 360
QY 1081 GGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 461 ArqArgLeuSerAspSerGlnTrpIleGluValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 481 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
QY 1201 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
DB 401 LeuAlaGlyCysGluGluArgSerAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 AFGCGTAATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1296
DB 421 lleValAsnGluAlaArgIleProAlaCysSerLeu 432
RESULT 7
Q8XC29
ID Q8XC29 PRELIMINARY: PRT: 434 AA.
AC Q8XC29
DT 01-MAR-2002 (TREMBLrel, 20, Created)
DT 01-MAR-2002 (TREMBLrel, 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel, 21, Last annotation update)
DE Phosphohydride phosphorylase, pH 2.5 acid phosphatase,
DE periplasmic.
GN APPA OR 21397 OR E3811d6.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RX SEQUENCE FROM N A
RX MEDLINE=0157:H7 / EDI933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RX Pose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RX Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RX Grotbeck K.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RX Apolara J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RX Welch P.A., Rattiner F.P.;
RT "Genome sequence of enterohemorrhagic Escherichia coli O157:H7";
RL Nature 406 529-534(2001);
RP SEQUENCE FROM N A
RC STRAIN=0157:H7 / RIMD 0509952;
RX MEDLINE 21156241; PubMed 11208796;
RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RX Han C.-G., Ohtsuki E., Nakayama K., Yurata T., Tanaka M., Tobe T.,
RX Iida T., Takai H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RX Kubara S., Shiba T., Hattori M., Shinofusa H.,
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K12";
RL DNA Res. 8:11-22(2001);
DR EMBL: AF005252; AAG5528.1; ALX157;
DR EMBL: AF002554; BAB34559.1;
DR InterPro: IPR000560; HisAc-phosphatase.
DR Pfam: PF00328; acid-phosphatase_1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
```

```
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Complete proteome.
SQ SEQUENCE 434 AA: 47337 MW: F197DF1869F9C4 CR664;
Alignment Scores:
Pred. No: 2 950-175 Length: 434
Score: 2144.60 Matches: 419
Percent Similarity: 96.54% Gaps: 0
Best Local Similarity: 96.54% Mismatches: 13
Query Match: 92.95% Indels: 2
DB: 16 Gaps: 1
US-09-866-379a-9 (1-1308) x Q8XC29 (1-434)
QY 1 ATCAAGCGATCTTAATCCCATTTTATGCTCTTGATTCGGTTAAGCGCGGAATTCGA 60
DB 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGT-----GAGCGGGAGCTCAAGCTGCGAAAGTGTGCTGATTCATCTGT 114
DB 21 PheAlaGlnSerGluProGluProGluLeuLysLeuGluSerValIleValSerArg 40
QY 115 CATGCTGTGGTGGTGGTAAATCAAGGCAATCAATCAATCAATCAATCAATCAATCAATCA 174
DB 41 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGluAspValThrProAspAla 60
QY 175 TCGCAACCTGGCGGTAAACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 234
DB 61 TrpLeuSerProValLysGlyTrpLeuThrProArgGlyGlyGlnLeuLeuLeuLeu 80
QY 235 TATCGCGCATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 294
DB 81 TyrLeuGlyHisGlyGlyGlnArgGlnArgLeuValAlaAspGlyLeuThrLysLysGly 100
QY 295 TCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 354
DB 101 CysProGlnProGlyGlnValAlaIleLeuAlaAspValAspValArgThrArgLysThr 120
QY 355 GCGCAACCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 414
DB 121 GlyGlnAlaPheAlaGlyGlyGlnAlaProAspGlyGlnThrValHisThrGlnAla 140
QY 415 GATACGTCATGTCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
DB 141 AspThrSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp 160
QY 475 AATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 534
DB 161 AsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThr 180
QY 535 GGCATTATCAAAACGGCGTTCGCGAAGTGGAACTGGAACGGGTGCTTAATTTCCGCAATCA 594
DB 181 GlyHisArgGlnThrAlaPheAlaGlyGlnLeuGluArgValLeuAsnPheProGlnSer 200
QY 595 TTCTGCTTAAAGCTCAGAAATAGGAGCGAAAGCTGTTCATTACGCGAGCAATTCATCG 654
DB 201 LeuCysLeuAsnArgCysLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSer 220
QY 655 CAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 714
DB 221 GluLeuLysValSerAlaAspAspValSerLeuThrGlyAlaValArgLeuAlaSerMet 240
QY 715 CTGACCGCATATTTCTCTCGGCAACAAATCAATCAATCAATCAATCAATCAATCAATCAAT 774
DB 241 LeuThrGlnThrPheLeuLeuGlnGlnAlaThrGlyMetProGluProGlyTrpLysArg 260
QY 775 ATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 834
DB 261 LeuThrAspSerHisGlnThrPheSerThrLeuLeuSerLeuHisAspAlaGlyPheTrpLeu 280
QY 835 CTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 894
DB 281 LeuGlnArgThrProGluValAlaArgSerAlaAlaThrProLeuLeuAspLeuIleMet 900
```


QY 1144 ATGAGTCAATAAAGCGCGTGGTATGATTAATAACAGCGCGGAGAGAGTGAACCTGACCCCG 1203
 DB 359 LeuArqAsuAlaAspAlaLeuThrLeuGlnAlaProAlaGlnArqValThrLeuGluLeu 488
 QY 1204 GTAGAGTCAAGAGCGGAAAGCGAGCGATCGTTGGTTGGCAGGTTTACGAAATG 1263
 DB 389 SerGlyCys---ProIleAspAlaAspGlyPheCysProMetAspLysPheAspSerVal 407
 QY 1264 GTGAATGAAGCA 1275
 DB 408 LeuAsuGlnAla 411
 RESULT 10
 Q827P1 PRELIMINARY: PKI: 413 AA.
 A* Q827P1:
 01-MAR-2002 (FEMBLrel_20, created)
 01-MAR-2002 (FEMBLrel_20, last sequence update)
 01-JUN-2002 (FEMBLrel_21, last annotation update)
 DE Glucose-1-phosphatase (G1Pase), secreted.
 GN SPY1153.
 OS Salmonella typhi.
 OW Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OW Salmonella.
 OX NCBI_taxid 601;
 RI SEQUENCE FROM N.A.
 RE STRAIN CT18.
 RX MEDLINE 21544947; PubMed-11677608;
 RA Parkhill J., Brown G., James K.D., Thomson N.R., Pickard D., White J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham E., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies K.M., Dowd L., Farrar J.,
 RA Freilich J., Hamlin N., Harper A., Hien T.T., Holtzky S., Jancels K.,
 RA Kiroh A., Latson T.S., Leather S., Meale S., O'Garra P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skellern J., Stevens K.,
 RA Whitehead S., Barrrell H.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18";
 RL Nature 413:848-852(2001);
 DR EMBL: AL627269; CAP08242.1;
 DR InterPro: IPR005660; HisAc-phosphatase.
 DR Pfam: PF00428; acid_phosphatase_1.
 DR PROSITE: PS00416; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 413 AA: 45630 MW: 9495636RHHKQV26 CRG64;
 Alignment Scores:
 Pred. No.: 2,36e-48 Length: 413
 Score: 547.00 Matches: 138
 Percent Similarity: 49.88% Conservative: 56
 Best Local Similarity: 33.74% Mismatches: 181
 Query Match: 23.69% Indels: 24
 DB: 16 Gaps: 10
 US-09-866-379a-9 (1-1308) x Q827P1 (1-413)
 QY 64 GTCACAGTCGAGCGGAG---CTCAAGTCGGAAGTCGTGATGTCACGTCCTCAGTGT 120
 DB 22 AlaGlnThrThrProGluGlyTyrGlnLeuGlnAlaValLeuMetMetSerArqHisAsn 41
 QY 121 GTGGTGTCTCA---ACCAAGGCAAGCAACTGATGTCAGATGTCACCGCAGACGATGG 177
 DB 42 LeuArqAlaProLeuAlaAsuAsuGlyAsuValLeuAlaGlnSerThrProAsuAlaTrp 61
 QY 178 CTAACCTGGACGTAAATCTGATGAGCTGCACACCGGCGCTGCTGACCTAATCGCTAT 247
 DB 62 ProAlaTrpAspValProGlyTyrGlnLeuMetThrThrLysGlyValLeuGluValTyr 81
 QY 238 CTGAGATATATACGGCTAGCGTCTGGTAGCGGCAATGTCGTGCTAAATGTCGCTGC 297
 DB 359 LeuArqAsuAlaAspAlaLeuThrLeuGlnAlaProAlaGlnArqValThrLeuGluLeu 488

DB 82 MetGlyPheIserThrArqGlnTrpLeuValAlaLeuGlyLeuThrProSerGlyLeuCys 101
 QY 298 CCGACGCTCGTCAAGTCGGGATTATTTGATGCTGCAAGCAAGCAATACCAATAAAGAGAG 463
 DB 102 ProAlaTrpAspThrValTyrAlaTyrAlaAsuSerLeuThrAlaThrValAlaThrAla 123
 QY 358 GAAGCTCTCGCGCGCGTGGCACTGATGCTGCAAAACCTGATACATACATACATACATACAT 419
 DB 122 GlnPheIleThrSerAlaPheProCysAspThrValHisHisGlnGlnCys 141
 QY 418 AGCTCAGTCGCGGATGTTTAAATCTCTAAACCTGAGGCTGCTTAAATCTGATAAAT 479
 DB 142 MetGlyThrMetAspProThrPheAsuProValThrAspAspGlnAlaAlaThrAla 161
 QY 478 GTGAATCTGATCAAGCGATCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 499
 DB 162 GlnGlnAlaValGlnAla---MetGlnLysAlaTrpSerGlnGln 180
 QY 538 CATTATCAAAACGGCGTTCGGAACTGGAAACGGCGCTTAAATCTGCGCAAAATCTG 199
 DB 176 HisLeuAspGluSerTyrLysLeuLeuGlnGlnThrHisThrLysGlnAspSerProSer 195
 QY 598 TGGCTTAAAGCTGAGAAATAGSAGGAAAGTCTTCAATAAGAGAGATTAATAGAGAGAGAG 609
 DB 196 Cys-----LysGlnLysHisThrCysSerLeuThrAspAlaLysAspThrPhe 111
 QY 658 CTCAACCTCAAGCGCACTCTCTCTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
 DB 212 SerAlaAsuTyrGlnGlnGlnThrGlyValGlnGlyThrLeuLysValGlyAsuSerLeu 141
 QY 718 AGCAGATATTTCTCTCTCAACAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 771
 DB 222 ValAspAlaPheThrLeuLeuTyrTyrLeuLeuPheThrMetGlnAspThrValAlaThrPhe 91
 QY 772 AGATACGATTTCAACACCTGAGAACTTCTGATTAAGATTTGATTAAGATTTGATTAAGAT 193
 DB 252 GlyLeuHisThrAspAlaPheThrLysValLeuSerLysLeuLysAsuLysTyrHisAsp 171
 QY 842 TTGCTACAGCAAGCGAGCTGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 891
 DB 272 SerLeuPheThrSerThrValAlaGlnAsuValAlaGlnAlaProLeuValLysTyrLeu 291
 QY 892 AAGACAGCGTTGACCGCCCATCCACCGCAAAACACAGCTGATGCTGCTGATTAATACAT 941
 DB 292 AspLysValLeuVal-----AlaGlnAlaGlnValSerAlaPro 909
 QY 952 TCAGTCGCTTATCGCGGACACGATACCTAACTGCTGCAATCTGCAAGAGAGAGAGAGAGAG 1011
 DB 304 LysValThrValLeuValGlyHisAspSerAsuThrAlaSerLeuThrAlaLeuAsp 123
 QY 1012 CTCAAC---TGGAGCTTCGCGGCTGACCGCGATACAGCGCGAGAGAGAGAGAGAGAGAGAG 1008
 DB 324 PheLysThrGlyTyrGlnLeuHisAspThrValGlnArqThrProThrLeuGlyGlnVal 143
 QY 1069 TTTCAACGCTCGCGCTCGCGCTAAAGGATAAAGATCAGTCTGATCTGATCTGCTGCTGCT 1128
 DB 344 PheGlnArqTrpHisAspGlyAsuAlaAsuArqAspLeuMetLysLeuGlyTyrValTyr 963
 QY 1129 CAGACTTTCAGCAGATGCTGATATAAAACCGCGCTGCTGATTAATAATACAGAGAGAGAGAG 1188
 DB 364 GlnSerAlaArqGlnLeuArqAsuAlaGlnAlaLeuThrLeuLysSerProAlaThrArq 483
 QY 1189 GTCAAACTGACCTCGCAGGATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1248
 DB 484 ValThrLeuGlnLeuLysGlyCysProValAspAlaAsuThrPheCysProLeuAsp 407
 QY 1249 GATTTCAGCAAAATCGTGAATGAAGAGAA 1275
 DB 403 LysPheAspAsuValMetAsuThrAla 411
 RESULT 11
 Q9AA04
 ID Q9AA04 PRELIMINARY: PKI: 414 AA.


```
QY 439 TTAAATGCTCTAAAAATGCGTTTTCGCAAC1GATATAACCGAAGCGGACGACGACGATC 498
Db 95 ----- 95
QY 499 CTCGAGAGCGACGAGAGGTCGATTTGATGTCGATTTACCGGCGATTTAAATGAAATGTTTCGC 558
Db 96 ----- 101
QY 559 GAATCGAAGCGGTCGTTAATTTTCGCGCAATCAAACTTGCGCTTAAACCTGACAAAATAG 618
Db 101 ----- 101
QY 619 GAGGAAAGCGCTTCATTAAACGACGCGATTTACCATGCGAATCAAGTCGACGCGGACTGT 678
Db 101 ----- 101
QY 679 GTCTCATTACCGCT---GCGGTAAAGCTCTGCAATCAATGCTGAGCGAGATATTTCCTCG 735
Db 102 LeuSerLeuValGlyGluProLeuArgAlaIlePheValLeuAspThrLeuMetCys 121
QY 736 CAAACAAGACAGGGAATCGCGGAGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 795
Db 122 GluAlaHisGlyLeuProLeuProAlaIleProAla-----SerPro 135
QY 796 AATACCTTCGTAATTTTCGATAAGCGGCAATTTTCATTTGCTGATACAAAGCGACGCGACAGGTT 855
Db 136 AspValLeuArgThrLeu-----AlaGlnIleSerAlaLeuAspTleGlyAlaHisVal 153
QY 856 GCGGAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 903
Db 154 GlyProArgAlaAlaGlnIleValGlnLeuThrGlyGlyIleLeuLeuAsnAlaIle 173
QY 904 AGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 963
Db 174 LeuAlaAsnPheSerArgValGlnArgLeuGly-----LeuProLeuLysMetValMet 191
QY 964 ATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1023
Db 192 TyrSerAlaHisAspSerThrLeuLeuAlaLeuGlnGlyAlaLeuGlyLeu----- 208
QY 1024 GTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
Db 209 - TyrAspGlyHisThrProProTyrAlaAlaCysLeuGlyPheGluPheArg 225
QY 1081 GGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1131
Db 226 CysHisLeuGlyAsnProAlaLysAspGlyGlyAsnValThrValSerLeuPheTyrArg 245
QY 1132 ACTTTACAGCAGATGCGGTGATAAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1191
Db 246 -----AsnAspSerAlaHisLeuProLeu----- 253
QY 1192 AAATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1251
Db 254 ProLeuSerLeuGlyCys-----TTCAlaProCysProLeuGlyArg 268
QY 1252 TTAAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1305
Db 269 PheTyrGlnLeuThrAlaProAlaAsnProProAlaHisGlyValSerCysHis 286
```

Search completed: December 13, 2002, 16:05:13
Job time : 105 secs


```

QY 640 CAGGATATACCAATGCAACATCAARHAGAGGACATCTGCTCTCATTAACAGGTCGGCTA 699
DB 245 ---LeuProGlyValThrLeuThrAspGluAspValValSerLeuMetAspMetCys 262
QY 700 AAGCTGTCATCAAGCTGACAGGATATATCTCTGCAACAAACACAGGAAATGCGGAG 759
DB 263 SerPheAspThrVal----- 267
QY 760 CAGAGGACAGCAAGATCAAGATATCAATATGATGGAACACATCTGTAAGTTG----- 813
DB 268 ---AlaArgThrSerAspAlaSerGlnLeuSerProPheCysGlnLeuPheThr 284
QY 814 CATAAC-----GGCAATTGATTGCTACAAACG----- 843
DB 295 HisAspGlnThrPheGlySerThrAsnTyrlsGlnSerThrGlyIysTyrlsGlyTyrlsGly 364
QY 844 -----AGGCAGAG-----GTTGCGGCG 861
DB 305 AlaGlyAspProLysGlyPheAlaLysGlyIleGlyPheThrAsnGlnLeuGlnAlaArg 324
QY 862 AGGAGGACAGCGCGTTATTAGATTGATCAAGACA-----GGTTGACGCGCCCATCCA 915
DB 325 LeuThrArgSerProValGlnAspHisThrSerThrAsnSerThrLeuValSerAsnPro 344
QY 916 CAGCAAAACACAGGATATGATGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 969
DB 345 -----AlaThrPheProLeuAsnAlaThrMetTyrValAspPhe 357
QY 970 GAGCAATATACTATCTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1024
DB 358 SerHisAspAsnSerMetValSerIlePhePheAlaLeu----- 370
QY 1030 GATCAAGCGATCAACACAGGCGGCTGCTGCAATCTCTTTTAACTGCTGCGCGGCTA 1089
DB 371 GlyLeuTyrAspGlyThrGlnProLeuSerArgThrSerValGlnLysAlaIysGlnLeu 399
QY 1090 AGTCAT---AACAGCAGCTGCAIT-----CAGGTTGGCGTGGCTCTCCAGACITTA 1137
DB 391 AspGlyTyrSerIleValValPheGlyAlaArgAlaTyrPheGlnThrMet 410
QY 1138 CAGCAATGCGTGATCAAAACGCGCTGCTCATTAATACGCGAGGAGAGAGTCAAAATG 1197
DB 411 GlnTyrLysSerGlyTyrGlyProLysGlyPheAlaValArgAlaIleAspAspArgValVal 1229
QY 1198 AATCTGACAGATGTCAGAGACGCAATATGCGAGGACATGCTTGTGTTGTCAGGTTTACG 1257
DB 430 ProLeuGlnGlyTyrAspValAspGlyLysGlyAlaGlySerGlnLysAlaValVal 448
QY 1258 CAAATGCAATCAAGCAACG 1278
DB 449 LysGlyLeuSerThrAlaArg 455

```

RESULT 10

```

US-09-044-718-80
Sequence 80, Application US/09044718
Patent No. 6,491,605

```

GENERAL INFORMATION:

```

APPLICANT: KOSTREWA, Dirk
APPLICANT: PASAMONTES, Luis
APPLICANT: TOMSCHY, Andrea
APPLICANT: VAN LOON, Adolphus
APPLICANT: WACHS, Kurt

```

TITLE OF INVENTION: MODIFIED PHYTASES

```

FILE REFERENCE: Modified Phytases

```

```

CURRENT APPLICATION NUMBER: US/09/044,718

```

```

CURRENT FILING DATE: 1998-03-19

```

```

PRIORITY APPLICATION NUMBER: EP 97810175.6

```

```

PRIORITY FILING DATE: 1997-03-25

```

```

NUMBER OF SEQ ID NOS: 82

```

```

SOFTWARE: Patent In Ver. 2.1

```

```

SEQ ID NO 80

```

```

LENGTH: 445

```

```

: TYPE: FRT
: ORGANISM: Aspergillus fumigatus
US-09-044-718-80
Alignment Scores:
Pct. No.: 0.000789 Length: 445
Score: 118.00 Matches: 101
Percent Similarity: 34.91% Conservations: 69
Best Local Similarity: 20.74% Mismatches: 107
Query Match: 5.11% Indels: 156
DB: 4 Gaps: 26

```

```

US-09-866-379a-9 (1-1308) x US-09-044-718-80 (1-465)

```

```

QY 19 GATATTTATGCTGCTCTCATTAACATCAARHAGAGGACATCTGCTCTCATTAACAGGTCGGCTA 28
DB 52 ProPheSerLeuGlnAspGluAspValValSerSerLysLeuProLys 60
QY 74 GAGTCAAGAGTGGAAAAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 135
DB 69 AspCysArgIleThrLeuValGlnValLeuSerAlaHisGlyAlaAlaIleThrProHis 98
QY 136 -----AAGGCGACGCACTGATGACATATCAATCAATCAATCAATCAATCAATCAAT 174
DB 89 SerLysSerLysTyrLysTyrLysTyrLysTyrLysTyrLysTyrLysTyrLysTyrLys 198
QY 175 -----TGGCAACCTGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 216
DB 109 LysGlyLysPheValPheLeuLysThrTyrAsnTyrThrLeuGlyAlaAspAspLeuThr 128
QY 211 CAGCGAGCTGGGAGCAATATGCGCAATATGCGCAATATGCGCAATATGCGCAATATGCGCA 264
DB 129 ProPheGlyGlnIleLeuValAspSerGlyIleGlyPheTyrIleAlaIleTyrLysAla 148
QY 265 GIAGGACAGGATATGCTGCTCATTAATGCTGCTCATTAATGCTGCTCATTAATGCTGCTCA 324
DB 149 LeuAlaArgSerValPro-----PheLeuArg 358
QY 325 GTGATCTGACAGGAGTATGCGTGAACAAACAGGCAAGCTGCGAGGAGGAGTGG 378
DB 159 AlaSerGlySerAspArgValIleAlaSerGlyLysPheIleGlyPheIleGlnIle 398
QY 379 -----TAACTTGAATGCTGCTCATTAATGCTGCTCATTAATGCTGCTCATTAATGCTG 429
DB 179 AlaLysLeuAlaAspProGlyAlaThrAsnAlaAlaAlaProAlaIleLeuValThrLeu 498
QY 403 CATAACAAAGCAATACAGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 462
DB 199 ProLysSerGlnThrPheAsn-----AsnThrLeuAspHisGlyVal 474
QY 463 TGTGCTGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 479
DB 213 CysThrLysPheGlnAlaSerGlnIleLeuLysAspGlyVal 495
QY 520 ATGCTGACATTAACAGGATATTAATCAAAAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 579
DB 226 AlaAlaAsnPheThrAlaLeuPheAlaProAspThrLeuArg 598
QY 580 TTTGCGCAATCAAACTTGTGCTGCTTAAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 639
DB 239 -----AlaArgAlaLeuLysHis 644
QY 640 CAGGATATACCAATGCAACATCAARHAGAGGACATCTGCTCTCATTAACAGGTCGGCTA 699
DB 245 ---LeuProGlyValThrLeuThrAspGluAspValValSerLeuMetAspMetCys 762
QY 700 AAGCTGTCATCAAGCTGACAGGATATATCTCTGCAACAAACACAGGAAATGCGGAG 759
DB 263 SerPheAspThrVal----- 267
QY 760 CAGAGGACAGCAAGATCAAGATATCAATATGATGGAACACATCTGTAAGTTG----- 813
DB 268 ---AlaArgThrSerAspAlaSerGlnLeuSerProPheCysGlnLeuPheThr 843
QY 814 CATAAC-----GGCAATTGATTGCTACAAACG----- 843
DB 295 HisAspGlnThrPheGlySerThrAsnTyrlsGlnSerThrGlyIysTyrlsGlyTyrlsGly 364
QY 844 -----AGGCAGAG-----GTTGCGGCG 861
DB 305 AlaGlyAspProLysGlyPheAlaLysGlyIleGlyPheThrAsnGlnLeuGlnAlaArg 324
QY 862 AGGAGGACAGCGCGTTATTAGATTGATCAAGACA-----GGTTGACGCGCCCATCCA 915
DB 325 LeuThrArgSerProValGlnAspHisThrSerThrAsnSerThrLeuValSerAsnPro 344
QY 916 CAGCAAAACACAGGATATGATGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 969
DB 345 -----AlaThrPheProLeuAsnAlaThrMetTyrValAspPhe 357
QY 970 GAGCAATATACTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1024
DB 358 SerHisAspAsnSerMetValSerIlePhePheAlaLeu----- 370
QY 1030 GATCAAGCGATCAACACAGGCGGCTGCTGCAATCTCTTTTAACTGCTGCGCGGCTA 1089
DB 371 GlyLeuTyrAspGlyThrGlnProLeuSerArgThrSerValGlnLysAlaIysGlnLeu 399
QY 1090 AGTCAT---AACAGCAGCTGCAIT-----CAGGTTGGCGTGGCTCTCCAGACITTA 1137
DB 391 AspGlyTyrSerIleValValPheGlyAlaArgAlaTyrPheGlnThrMet 410
QY 1138 CAGCAATGCGTGATCAAAACGCGCTGCTCATTAATACGCGAGGAGAGAGTCAAAATG 1197
DB 411 GlnTyrLysSerGlyTyrGlyProLysGlyPheAlaValArgAlaIleAspAspArgValVal 1229
QY 1198 AATCTGACAGATGTCAGAGACGCAATATGCGAGGACATGCTTGTGTTGTCAGGTTTACG 1257
DB 430 ProLeuGlnGlyTyrAspValAspGlyLysGlyAlaGlySerGlnLysAlaValVal 448
QY 1258 CAAATGCAATCAAGCAACG 1278
DB 449 LysGlyLeuSerThrAlaArg 455

```


db 421 Cys-ProLeuAlaAsnTyrThrSerLeuLeuAsnLys 43.

Search completed: December 13, 2002, 16:07:40
Job time : 34.5 secs

```
db 170 .....SerSerGlyTyrGlyArgValIle .....Glu 178
QY 340 TGGTAACTGTAAGAGAGAGG---TTCGCGAGAGAGCTGGAGCTGACGTGGCAATA 396
db 179 ThrAlaArgLysPheGlyLeuGlyPheGlyTyrAsnTyrSerThrAsnAlaAlaLeu 198
QY 397 ACCATACATACCGACGATAGATAGCTGACGTCGACGCGCTTATTTAAATGCTCTAAAAACT 456
db 199 AsnIleIleSerGluSerGluValMetGlyAlaAspSerLeu-----Thr 213
QY 457 GAGGTTCGCTAACCTGATAAAGGAACGAGCTGACGCGCATGCTGACAGAGGCGCAGCAAG 516
db 214 ProThrCysAspThrAspAsnAspGlnThrThrCysAspAsnLeuThr----- 229
QY 517 TCAATTCGTACTTTAGCGGCGCATATATCAAAAGCGGTTTCGCGAATGGAACGGGTGCTT 576
db 240 -----TyrGln----- 241
QY 577 AATTTTCGTAATCAAACTTCGGGCTTAAAGCT---GAGAAACAGACAGCAAAAGCTGTTC 633
db 242 ---LeuProGlnPheLysValAlaAlaAlaLeuAsnSerGlnAspProGlyMetAsn 250
QY 634 TTAAGCAGGCACTATCAATCAACTCAAGCTGACGCGGACTGTGTCTCATTAACCGGT 693
db 251 LeuThr-----AlaSerAspValTyrAsnLeu----- 259
QY 694 GAGTAAACCTGCTGATCAATCTACAGGAAATATTTCTGCTGTAAATAAGCAGAGGAATG 753
db 260 ---MetValMetAlaSerPhe-----GluLeuAsnAlaAla 270
QY 754 CAGGAGACGGGCTGGGAAAGATACAGCGATTCACACAGTGGCAACAGCTTG----- 804
db 271 ProPheSerAsnTrrIleAsnAlaPheThrGlnAspGluTrrPvalSerPheGlyTyrVal 290
QY 805 -----CTAAGTTTCGAT----- 816
db 291 GluAspLeuAsnTyrTyrCysAlaGlyProGlyAspLysAsnMetAlaAlaValGly 310
QY 817 ---AAGGCAATTTCATTTGGCTATCAACGACGCGCACAGATTTCGCGCAGC 864
db 311 AlaValTyrAlaAsnAlaSerLeuThrLeuLeuAsnGlnGlyProLysGluAla----- 428
QY 865 CAGGCAACCGGTTATAGATTTCATCAAGACAGCTTCACGCGCGCATCCACCGCAAAA 924
db 328 ----- 328
QY 925 CAGGCTATGCTGCATATACCGACTTCAGTGTGTTATCGCGCGGACACGATACTAA 984
db 329 -----GlySerLeuThrPheAsnPheAlaHisAspThrAsn 340
QY 985 CTGCAAAATCTCGCGGCGTATCGAGCTC-----AAGTGCACGCTTCGC----- 1029
db 341 IleThrProIleLeuAlaAlaLeuGlyValLeuIleProAsnGlnAspLeuProLeuAsp 360
QY 1030 -----GTCAGCGG-----GATAACACGCGCGGACGCTGGTCAACTG 1065
db 361 ArgValAlaPheGlyAsnProTyrSerLeuLysAlaValProMetGlyGlyHisLeu 380
QY 1066 GCGTTCAAGGCTGC-----GTCGCGCTAACGATACACAGCAGCGGATTCAG 1113
db 381 ThrIleGluIleLeuSerCysGlnAlaThrAlaLeuSerAspGlnGlyThrTyrValArg 400
QY 1114 GTTTCGTGCTTCTTCAGATTTACAGCATATGCTGATAAAGAGCGCTGTCATTAAAT 1173
db 401 LeuValLeu----- 403
QY 1174 ACAGCGCGGACAGGTCAGAAATGACCTGTACATGCTGAGAGAGCAAAATGCGCAGGC 1233
db 404 ---AsnIleAlaValLeuProPheAsnAspTyrThrSerClyProGlyTyrSer 420
QY 1234 ATGTTGCTGGCGACGTTTACGCAAAATGTCAGAA 1272
db 1272 ----- 1272
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: December 13, 2002, 15:55:21, Search time 62 Seconds
(without alignments)
5622.316 Million cell updates/sec
Title: US-09-866-379A-9
Perfect score: 2309
Sequence: 1 atgaagcgatcttaatcc..... gaggatggagctctatcta 1308

Scoring table: BLOSUM62
Gapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum hit seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL frame_n2p.model -DEV xlp
-Q=us-09-866-379a-9 -FASTA=us-09-866-379a-9.fasta_1 1479
-DH=A.Geneseq_101002 -OFMT=fastan -SUFFIX=traq -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -STAR=1 -END=1 -MATRX=BLOSUM62 -TRANS=human40.cd:
-LIST=45 -DOCATL=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -GTRM=pcr -NORM=ext -HEAFISE=500 -MINI=FN=0 -MAXI=FN=200000000
-USER=US09866379A -CGN_1_1_45 -runat_13122052_144535_21075 -NCT=6 -ICP=3
-NO_XLPXY -NO_MAP -LARGEFQF -NEG_SCORES=0 -WAIT -LONGLOG DEV_TIMEOUT=126
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT 0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: A_Geneseq_101002:
1: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
2: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
3: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
4: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
5: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
6: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
7: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
8: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
9: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
10: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
11: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
12: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
13: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
14: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
15: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
16: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
17: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
18: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
19: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
20: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
21: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
22: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002
23: 51182 101002 101002 101002 101002 101002 101002 101002 101002 101002

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Length | DB ID | Description |
|------------|-------|---------|--------|-------|-------------|
| 1 | 2228 | 96.5 | 430 | 23 | AAE15808 |
| 2 | 2188 | 94.8 | 410 | 22 | AAH37872 |
| 3 | 2188 | 94.8 | 440 | 23 | AAE22836 |
| 4 | 2188 | 94.8 | 440 | 23 | AAE15806 |
| 5 | 2182 | 94.5 | 432 | 21 | AAU77775 |
| 6 | 2182 | 94.5 | 432 | 21 | AAH36257 |
| 7 | 2182 | 94.5 | 432 | 21 | AAH36259 |
| 8 | 2182 | 94.5 | 432 | 21 | AAH36259 |
| 9 | 2182 | 94.5 | 432 | 21 | AAH36259 |
| 10 | 2182 | 94.5 | 432 | 21 | AAH36259 |
| 11 | 2182 | 94.5 | 432 | 21 | AAH36259 |
| 12 | 2182 | 94.5 | 432 | 21 | AAH36259 |
| 13 | 2182 | 94.5 | 432 | 23 | AAH15807 |
| 14 | 2177 | 94.3 | 432 | 22 | AAH36261 |
| 15 | 2171 | 94.0 | 432 | 22 | AAH36261 |
| 16 | 2170 | 94.0 | 432 | 22 | AAH36261 |
| 17 | 2157 | 93.4 | 432 | 22 | AAH36261 |
| 18 | 2147 | 93.0 | 432 | 21 | AAH36261 |
| 19 | 2080 | 90.1 | 423 | 20 | AAH36261 |
| 20 | 1192 | 51.6 | 261 | 22 | AAH36261 |
| 21 | 676 | 26.4 | 140 | 22 | AAH36261 |
| 22 | 637 | 27.6 | 144 | 22 | AAH36261 |
| 23 | 590 | 25.5 | 123 | 22 | AAH36261 |
| 24 | 324 | 14.1 | 118 | 22 | AAH36261 |
| 25 | 263 | 11.4 | 1303 | 22 | AAH36261 |
| 26 | 201 | 8.7 | 426 | 22 | AAH36261 |
| 27 | 201 | 8.7 | 426 | 23 | AAH36261 |
| 28 | 201 | 8.7 | 426 | 23 | AAH36261 |
| 29 | 184 | 8.0 | 1062 | 22 | AAH36261 |
| 30 | 131 | 5.7 | 423 | 23 | AAH36261 |
| 31 | 127 | 5.5 | 395 | 22 | AAH36261 |
| 32 | 126 | 5.5 | 440 | 21 | AAH36261 |
| 33 | 126 | 5.5 | 465 | 19 | AAH36261 |
| 34 | 126 | 5.5 | 465 | 19 | AAH36261 |
| 35 | 126 | 5.5 | 465 | 19 | AAH36261 |
| 36 | 126 | 5.5 | 465 | 19 | AAH36261 |
| 37 | 126 | 5.5 | 465 | 19 | AAH36261 |
| 38 | 124 | 5.4 | 467 | 22 | AAH36261 |
| 39 | 124 | 5.4 | 467 | 21 | AAH36261 |
| 40 | 124 | 5.4 | 467 | 21 | AAH36261 |
| 41 | 124 | 5.4 | 467 | 21 | AAH36261 |
| 42 | 124 | 5.4 | 467 | 21 | AAH36261 |
| 43 | 124 | 5.4 | 467 | 21 | AAH36261 |
| 44 | 124 | 5.4 | 467 | 21 | AAH36261 |
| 45 | 118 | 5.1 | 440 | 21 | AAH36261 |

ALIGNMENTS

RESULT 1
AAE15808
ID AAE15808 Standard, Protein: 430 AA.
XX AAE15808:
AC
UU 25-MAR-2002 (first entry)
XX Escherichia coli appA phytase mutant protein.
XX Bacterial phytase, K12 appA phytase, protease substrate; anabolic;
KW gas count-stimulant; nutrient; vitamin; food treatment process; therapy;
KW chemical tolerance; growth performance; alcoholic drink; biopulping;
KW non-alcoholic drink; biobleaching; mutant; mutagen.
XX Escherichia coli.
OS Synthetic.
XX
FH Key Location/Qualifiers


```

Db      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluAraTrp 360
QY      1081 GATGGGCTTAAGGATAAACAGGACAGTGGATTCACCTTTGGCTGCTGCTCCACACTTAAAG 1140
Db      361 AraAraLeuSerAspAsnSerGlnTrpPheGluValSerLeuValPheGlnThrLeuGln 480
QY      1141 CAGATGCGGATAAAGGCGGTGGTCATTAAATATACGCGCGGACAGGAGGCAACTGAC 1200
Db      481 GlnMetAraAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 490
QY      1201 CTGGACGATATGAAACAGGCAATATGCTAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db      401 LeuAlaGlyCysGlnArgAsnAlaGlnIleGlyMetCysSerLeuAlaGlyPheThrLeu 420
QY      1261 ATCTGCAATCAACACGACATACGCGGTGGTGCATTTCAATATTCAT 1305
Db      421 ThrValAsnGluAlaAraGlnIleProAlaCysSerLeuAraSerHis 435
RESULT 5
AA077775 standard; Protein; 432 AA.
XX
AA077775.
XX
05-JUN-2002 (first entry)
XX
Phytase associated protein.
XX
Phytase.
XX
Gardner et al.
XX
KR99096028-A.
XX
15-DEC-1999.
XX
25-MAY-1998: 98KR 0018810.
XX
25-MAY-1998: 98KR-0018810.
XX
(WOJ-) WOJIN CO LTD.
XX
Bae HD, Forceburgh CW, Goloben S, Chong RJ.
XX
WPI: 2000 645078/62.
XX
N-DBS; ABK12514.
XX
Novel phytase gene, recombinant phytase and usage thereof -
XX
disclosure; Fig 3; 10pp; Korea.
XX
The invention relates to a novel phytase gene, a recombinant
XX
phytase gene and their uses, this is the amino acid sequence of the
XX
phytase associated protein described in the invention.
XX
Sequence 432 AA.
XX
Alignment Scores:
XX
Pred. Res. 254e 221 Length 432
XX
Score: 2182.00 Matches: 424
XX
Percent Similarity: 98.15% Conservative: 0
XX
Best local Similarity: 98.15% Mismatches: 8
XX
Query Match: 94.50% Indels: 0
XX
DB: 21 Gaps: 0
XX
US 09-866-379A-9 (1-1408) x AA077775 (1-432)
QY      1 ATGAAGCGATGATTAATCGCATTTTAATGCTTTGATTCGTTAAGCCGCAATCTGCA 60
Db      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY      61 TCGGCTACAGAGACGCGACACTGAAGTGGAAATGGAAATGTGGTGAATGTCATGTCATGCT 120

```

```

Db      21 PheAlaGlnSerGluProGluLeuLysLeuIleSerValIleValIleValIleValIle 40
QY      121 GAGAGGCTTCAACCAAAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 180
Db      41 ValAraGlnProThrLysAlaThrThrGlnLeuMetGlnAspValThrProAspAlaTrpLeu 400
QY      181 AACCTGGCGGGTAAACCTGGGCTGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240
Db      61 ThrTrpProValLysLysLeuGlyTrpPheThrProGlnGlyGlnLeuLeuAlaTrpLeu 80
QY      241 GAACTATTAGTGGCGGCTAGGCTGGTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 300
Db      81 GlyHisThrGlnAraGlnAraLeuValAlaAspGlyLeuLeuValLysLysLysPro 100
QY      301 CAGCTGGTCAAGTGGCAATTAATGCTGATGCGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360
Db      101 GlnSerGlyGlnValAlaIleIleValAspValAspGlnAraThrAraLysThrGlyGln 120
QY      361 GCGTTACCGCGCGGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
Db      121 AlaPheAlaAlaGlyLeuAlaIleProAspLysAlaIleIleValIleValIleValIle 140
QY      421 TCGAGTGGCGATTCGGTGAATTAATGCTGATGCGAAGCAAGCAAGCAAGCAAGCAAGCAAG 480
Db      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY      481 AACGTCATTCAGTGGCAATTCAGTGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 540
Db      161 AsnValThrAspAlaIleLeuSerAraAlaGlyLysSerIleAlaAspIleThrGlyHis 180
QY      541 TATCAAAAGGCGCTTCGCGAAGTGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 600
Db      181 ArgGlnThrAlaPheAraGlnLeuGlnAlaValLeuAspPheThrLeuSerAsnLeuGlyS 200
QY      601 CTTAACAGTGAGCAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 660
Db      201 LeuLysAraGlyLysGlnAspLeuSerCysSerLeuThrGlnAlaIleProSerLeuGln 220
QY      661 AAGCTGACGCGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db      221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 740
QY      721 CAGATAATTTCTCTGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 780
Db      241 GluIleThrLeuLeuGlnIleAlaGlnIleGlyMetProIleProGlyIlePheGlyAlaThr 800
QY      781 GATTCAACCTGCTGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 840
Db      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuGlnAsnAlaGlnIleProSerLeuGln 900
QY      841 CAGTGGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
Db      281 ArgThrProGluValAlaAraSerAraAlaThrProLeuLeuAspLeuLysLysLysAla 1000
QY      901 TTAGAGGCGCATCCACGCAAAAAAAGAGAGATATGCTGCAATTCATTTCTTTAAACAAGA 1060
Db      301 LeuThrProHisProGlnLysGlnAlaGlyValThrLeuThrThrSerValLeuGln 1080
QY      961 TTAAGCGCGCAACAGATTAATGCTGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1140
Db      421 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyLysAlaLeuPheGlnAsnAla 1160
QY      1021 AAGCTTCGCGCTGACGCGCGATAAACAGGCGCTAGGCTGCAAGTGGCTGCTGCAAGTGGCT 1080
Db      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlnAraTrp 400
QY      1081 GCGAGCTAAGAGCAAAACAGGCAAGGCAATGAGTGGATTCAGTGGATTCAGTGGATTCAGT 1140
Db      361 AraAraLeuSerAspAsnSerGlnTrpPheGlnValSerLeuValPheGlnThrLeuGln 480
QY      1141 CAGATGCGGATAAAGGCGGTGGTCATTAAATATACGCGCGGACAGGAGGCAACTGAC 1200

```


Db 401 LeuAlaGlyTyrSerGluAsnAlaLalleLeuMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 APSTGTAATTAAGGACGACATACCGGGGACACTTC 1296
 Db 421 TleValAsnGluAlaArqIleProAlaCysSerLeu 432
 RESULT 10
 ID AAB36261 standard; Protein: 432 AA.
 XX AAB36261;
 XX 20-FEB-2001 (first entry)
 XX R15/APPA plasmid translated sequence.
 DE Transgenic animal; salivary protein; phytase; phosphorus, animal growth;
 KW environmental pollution; pig.
 XX Rattus sp - chimeric.
 OS Escherichia coli - chimeric.
 XX W0200064247-A1.
 IN 02-NOV-2000.
 XX 20-APR-2000; 2000W0-CAC0430.
 PF 24-APR-1999; 99DS-0106509.
 XX (YIGH-) UNIV GUELPH.
 XX Forsberg CW, Golevan S, Phillips JP;
 DR WPL: 2000-687245/67.
 DR N-PSDB: AAC68298.
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein; protein; phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX Disclousure; Fig 21; 152pp; English.
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland, low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence.
 XX Sequence 432 AA:
 SQ
 Alignment Scores:
 Pred. No.: 2,540-221 Length: 432
 Score: 2192.00 Matches: 424
 Percent Similarity: 98.15% Conservative: 0
 Best Local Similarity: 98.15% Mismatches: 8
 Query Match: 94.50% Indels: 3
 DB: 21 Gaps: 0
 US-09-866-379A-9 (1-1308) x AAB36261 (1-432)
 QY 1 ATGAAGAGATCTTAAGCCGATTTTATGCTTCTGATTCGTTAATCCCGCAATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAAGTGAAGCGGAGACTGAAGCTGGAAGTGTGGAGATTCACATCGTCATGCT 120
 Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
 QY 121 GTGATGCTGCAACCAAGGTCACAGCACTGATGACGATGATGACGACGATGACGAC 180

Db 41 ValAlaGluIleThrLysAlaIleGlnLeuMetGlnAspValIleThrProAspAlaIlePro 60
 QY 181 ACCTGCTGCTTAAATTTGGGTCGAGCTGACATAGAGAGAGGTTGGAGACTAATACCTTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTyrIleThrProArgGlyGlyGluLeuIleValAlaIleLeu 80
 QY 241 GGACATTACTGGCGTCACGCTCTGTAAGCGACGATTCGCTGCTTAAATGATGATGATGCG 300
 Db 81 GlyHisTyrGlnArgGlnArgGlnValAlaGlyLeuValAlaGlyLeuValAlaGlySerPro 100
 QY 301 CAGCTTGTGTCACGCTGATTCGTCACGACGACGACGACGACGACGACGACGACGACGAC 360
 Db 101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspGlnArgThrArgLysThrGlyGln 120
 QY 361 GTCTTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
 Db 121 AlaThrAlaGlyLeuAlaIleAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCCAGTCCGATGCTGCTTAAATTTGGGTCGAGCTGACATAGAGAGGTCGACATTCACGCGCAT 540
 Db 141 SerSetProAspProLeuPheAspProGlnLysThrGlyValCysGlnLeuAspAspAla 160
 QY 481 AAATGATGACGATCTGACAGAGGTCAGAGGTCGACATTCGTCACATTCACGCGCAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 QY 541 TATCAAACTGGGTCGAGCTGTAAGCTTAAATTTGGGTCGACATTCGTCACATTCGTCAC 600
 Db 181 ArgGlnThrAlaPheArgGluGluArgValLeuAspPheProGlnSerAspAlaCys 200
 QY 601 CTAAAGGTCAGAGGTCGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 660
 Db 201 LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
 QY 661 AAATGATGACGATCTGACAGAGGTCAGAGGTCGACATTCGTCACATTCACGCGCAT 720
 Db 221 LysValSerAlaAspAspValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuAla 240
 QY 721 CACATATTTCTTCTGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 780
 Db 241 GlnIlePheLeuLeuGlnAlaGlnGlyMetProGlnProGlyTyrProGlyArgIleThr 260
 QY 781 GATTCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 840
 Db 261 AspSerLeuGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaCLeuPheTyrLeuLeuGln 280
 QY 841 GGTATGTCACAGGTTGCTGTCACGACGACGACGACGACGACGACGACGACGACGACGAC 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleCysThrAla 300
 QY 901 TTCAGGTCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 960
 Db 301 LeuThrProGlnSerProGlnSerLeuArgGlyValThrAlaProSerValLeuGln 320
 QY 961 TTATGCTGCTTAAATTTGGGTCGAGCTGACATAGAGAGGTCGACATTCACGCGCAT 1020
 Db 321 PheIleAlaGlyHisAspThrAsnAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 AATGCTGCTTAAATTTGGGTCGAGCTGACATAGAGAGGTCGACATTCACGCGCAT 1080
 Db 341 ThrLeuProGlyLeuProAspAsnThrProProGlyValLeuLeuValPheGlnSerTrp 360
 QY 1081 GTGCTGCTTAAATTTGGGTCGAGCTGACATAGAGAGGTCGACATTCACGCGCAT 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATGCTGATAAATGCGCTGCTCATTAATATAGCGCGGCGGAGAGAGTGAATATGAC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGCGCAGATGTAAGAGCGCAATGCGACGACGACGACGACGACGACGACGACGACGAC 1260
 Db 401 LeuAlaGlyTyrSerGlnArgAlaSerAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420

QY 1261 ATGCGAATGAAACACGATATATGAGGTCACATTC 1296
 Db 421 TLeValAsnGluAlaArgTLeProAlaGlySerLeu 432

RESUME 12

AA046263
 ID AAB46263 standard; Protein: 432 AA.

XX
 AC AAB46263;

DT 20-FEB-2001 (first entry)

XX lamaz/APPA plasmid translated sequence

XX Transgenic animal; substrate product, phytase, phosphatase, animal growth,
 KW environmental pollution; pig.

XX Mos musculus - chimeric.

OS Escherichia coli - chimeric.

XX W0200064247-A1.

PN 02-NOV-2000.

XX 20-APR-2000; 2000W0-CA000430.

PR 23 APR-1999; 99DS-0130508.

PA (U93)- UNIV GUELPH.

XX Forsberg CW, Golembo S, Phillips JP;

XX WPI; 2000-087245/67

DR N-PSDB; AAC68300.

XX Transgenic non-human animal for gastrointestinal tract specific
 PI expression of a protein, preferably phytase, comprises a nucleic acid
 PI sequence including a heterologous transgene construct encoding the
 PI protein -

XX Dislosures; Fig 23, 152pp; English.

XX The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland, low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence.

XX Sequence 432 AA:

Alignment Scores:
 Prod. No.: 2540-231 Length: 432
 Score: 2182.00 Matches: 424
 Percent Similarity: 98.15% Conservatives: 0
 Best Local Similarity: 98.15% Mismatches: 8
 Query Match: 94/50% Indels: 0
 DB: 21 Gaps: 0

US-09-866-379A-9 (1-1308) x AAB36263 (1-432)

QY 1 ATGAAGCGATCTTAATCCGATTTTATGCTCTCTGATTCGTTAAACCGGCAATCTGCA 60

Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

QY 61 TTGGCTCAGACTAGCGGAGCTCAAGCTCGCAAACTGCTGCTGATTCAGTCGTCATGGT 120

Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40

QY 121 GTGGTCTCTGCAACGAGGCGACGCAATGATGCTAGTATGACCCGACAGCGAATGCCA 180

Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrPheAspAlaThrPhe 60

QY 181 ACCTGGCGCGCTAAACGCGCGGTCACGTCACACGCGCGCGGCTGCTACGTAATGCGTAATG 240
 Db 61 ThrTrpProValLysLeuGlyTyrPheLeuThrProArgGlyGlyGlyGlyGlyGlyGlyGly 80
 QY 241 GGAATTATTGCGGTGAGGCTGCTAGCGGAGGAGGATGCTGCTGCTGCTGCTGCTGCTG 400
 Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspTyrLeuLeuAlaTyrLysGlyGlyGlyGly 100
 QY 301 CACTCTGCTCAGCTCGCGATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 460
 Db 101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspPheLeuThrArgLysThrGlyGly 120
 QY 361 GCTTCTGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 121 AlaIleAlaGlyLeuAlaIleAspLysAlaIleLeuValAspGlnGlnAlaAspThr 140
 QY 421 TCCAGTCCCGATCGTTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 Db 141 SerSerProAspProLeuPheAspThrLysThrGlyValLysGlnLeuAspAsnAla 160
 QY 481 AACGTACTCAGCGCGATCTGCGAGCGCGAGGGTCAATTGCTGCTGCTGCTGCTGCTGCT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyLysLeuAlaAspPheThrGlyHis 180
 QY 541 TATCAAAAGCGGTTTCGCGAACTGGAAGCGGTGCTTAATTTTCGCAATTAATTTGCTG 600
 Db 181 ArgGlnThrAlaPheArgGlnLeuGluArgValLeuAsnProGlnSerAsnLeuLys 200
 QY 601 CTTAACTGAGAAACAGACGAAAGCTGTTCATTAGCGGAGGATTATTAACGGAATC 660
 Db 201 LeuLysArgGlyLysGlnAspGluSerLysGlySerLeuLeuAlaLeuPheGlnLeu 220
 QY 661 AAGCGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuValSerLeuVal 240
 QY 721 GAGATAITTCGCTGCAACAAAGCAAGAGAAATGCGGAGCGGAGGAGGAGGAGGAGGAG 780
 Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTyrPheArgIleThr 260
 QY 781 GATTCACAGAGAGAAACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 261 AsnSerHisSerLysPheSerLeuLeuSerLeuLeuSerLeuLeuSerLeuLeuSer 280
 QY 841 GAG 900
 Db 281 ArgThrProGlnValAlaArgSerArgAlaIleProLeuLeuLeuLeuLeuLeuLeu 300
 QY 901 TTGACGCGCGCTACCGCGCAAAACAGCGGTATGCTGAGATTACCACTTACAGCGCG 960
 Db 301 LeuThrProHisProProGlnLysGlnAlaTyrLeuValThrLeuPheThrSerValLeu 320
 QY 961 TTATTCGCGAGAGATATATATATCTGCAATCTGCGGAGGAGGAGGAGGAGGAGGAG 1020
 Db 321 PheIleAlaGlyHisAspGlnAspLeuAlaGlnLeuLeuLeuLeuLeuLeuLeuLeu 340
 QY 1021 AAGCTTCGCGAG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGluArgThr 360
 QY 1081 GCGCGGCTAAGCGATACACGCGAGCTGATTCAGGCTTCGCTGCTGCTGCTGCTGCTGCT 1140
 Db 361 ArgArgLeuSerAspAsnSerGluTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATGCTGATAAACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnValLysLeuThr 400
 QY 1201 CTGCGAGCATCTAAGACGCAAAATGCGGAGCGCTATGCTGCTGCTGCTGCTGCTGCTGCT 1260
 Db 401 LeuAlaGlyCysGluGlnArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 AATGTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1296

|||||
421 lleValAsnGluAlaArqIleProAlaCysSerLeu 432

RESULT 13

AAE15807

ID AAE15807 standard; Protein: 432 AA.

XX AC AAE15807;

XX XX

DT 26-MAR-2002 (first entry)

DE DE

XX Escherichia coli appA phytase wild type protein.

XX Bacterial phytase; K12 appA phytase; protease stability; anabolic;

KW gastro-intestinal; nutritive; feed treatment; process; therapy;

KW thermal tolerance; growth performance; alcoholic drink; biopulping;

KW non-alcoholic drink; tri-bleaching

XX XX

OS Escherichia coli.

XX XX

PR W0200190333-A2.

PR 29-NOV-2001.

XX XX

XX 24 MAY 2001. 2301W3 0S17118.

XX 25-MAY-2000; 200003 0580515.

XX XX

PA (DIVE-) DIVERSA CORP.

XX XX

XX Short JM, Kretz KA, Gray KA, Barton NF, Garrett JB, O'Donoghue E;

XX N PSDB; AA025463.

XX XX

PT New bacterial phytase for e.g. improving the nutritional value of
PT phytate containing foodstuffs and subsequently improving the growth
PT performance of an organism that consumes it, or in treating animal
PT digestive systems -
XX XX

PS Claim 54; Fig 8; 170pp; English.

XX XX

CC The patent discloses recombinant bacterial phytase from *Escherichia coli*
CC K12 appA phytase. The enzyme has phytase activity and improved thermal
CC tolerance when compared with wild-type phytase. It has improved process
CC stability at low pH. The recombinant phytase is useful for improving the
CC nutritional value of phytate-containing foodstuffs and subsequently
CC improving the growth performance of an organism that consumes it, in
CC treating animal digestive systems, in feed treatment processes and for
CC in vitro purposes related to research, discovery and development. They
CC are also used for generating recombinant digestive system life forms,
CC for producing or manufacturing alcoholic and non-alcoholic drinks based
CC on the use of moulds, grains and/or plants, in biopulping and bio-
CC bleaching where a reduction in the use of environmentally harmful
CC chemicals that are traditionally used in the pulp and paper industry
CC is desired and in the reduction or possible elimination of the need
CC for mineral supplements, enzymes or therapeutic drugs for animals
CC from the daily feed thus increasing the amount calories and nutrients
CC present in the feed. The present sequence is *E. coli* appA phytase
CC wild type protein.
XX XX

SQ Sequence 432 AA;

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|-----|
| Prod. No.: | 2 54e 221 | Length: | 432 |
| Score: | 2182.00 | Matches: | 424 |
| Percent Similarity: | 68.1% | Conservative: | 0 |
| Best Local Similarity: | 98.15% | Mismatches: | 8 |
| Query Match: | 94.50% | Indels: | 0 |
| DB: | 23 | Gaps: | 0 |

US-09-866-379A-9 (1-1308) x AAE15807 (1-432)

| | | | |
|----|------|---|------|
| QY | 1 | ATGAAAGGATCTTAATCCCAATATATATCTCTCTGATATGAGTAAATGCTAAATCTCA | 60 |
| DB | 3 | MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProIleProIleProIleProIle | 20 |
| QY | 61 | TTGCTCAGAGTGGCGGAGTGAATCTGCAAAAGTCTCTGATCTGCTACCTCTCAAGG | 120 |
| DB | 21 | PheAlaGlnSerGluProGlnLeuLysLeuGlnSerValValIleValSerAlaHisSer | 40 |
| QY | 121 | GTGGGCTCTCCAAACAGAGGCAAGCAATGATGATAGATGATGATGATGATGATGATG | 180 |
| DB | 41 | ValArgAlaProThrIleValIleThrGlnLeuMetGlnAspValIleProAspAlaIlePro | 60 |
| QY | 181 | ACCTGGAGGCTAAACCTGGAGTGAATGCAATGCAATGCAATGCAATGCAATGCAATG | 240 |
| DB | 61 | ThrTrpIleValLysLeuGlyTrpPheLeuThrPheAlaGlyGlnLeuIleValIleLeu | 80 |
| QY | 241 | GGAATATATCTGGGCTCAGCGCTGCTGATGATGATGATGATGATGATGATGATGATG | 300 |
| DB | 81 | GlyHisIleIleValIleValIleValIleValIleValIleValIleValIleValIle | 100 |
| QY | 301 | CAGCTGCTCTCAGGCTCGGCAATATGCTGATGATGATGATGATGATGATGATGATG | 360 |
| DB | 101 | GlnSerGlyGlnValAlaIleIleAlaAspValAspIleArgIleArgIleArgIle | 120 |
| QY | 361 | GGCTTGGAGGCGGCTGGAGTGAATGATGATGATGATGATGATGATGATGATGATG | 420 |
| DB | 121 | AlaPheAlaAlaGlyLeuAlaIleAspIleValIleValIleValIleValIleValIle | 140 |
| QY | 421 | TCAGTGGGATCTGCTGATATATATGCTGATGATGATGATGATGATGATGATGATG | 480 |
| DB | 141 | SerSerProAspProLeuPheAspProLeuLysThrGlyValPysGlnLeuAspAspAla | 160 |
| QY | 481 | AACTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG | 540 |
| DB | 161 | AsnValThrAspAlaIleLeuSerArgAlaGlyLysSerIleAlaAspPheThrGlyHis | 180 |
| QY | 541 | TAATCAAAAGGCTTCTAGCAACCTGATGATGATGATGATGATGATGATGATGATG | 600 |
| DB | 181 | ArgGlnThrAlaPheArgIleLeuGlnValIleAsnIleProIleSerAlaSerLeu | 200 |
| QY | 601 | CTTAAAGCTGAGAAACAGCAGCAGAAAGCTCTCATTAATGATGATGATGATGATG | 660 |
| DB | 201 | LeuLysArgGlnLysGlnAspIleSerLysSerLeuIleGlnAlaLeuProSerLeuLeu | 220 |
| QY | 661 | AAGCTGAGCGCGCACTCTGCTGATATAACTGCTGCTGATGATGATGATGATGATG | 720 |
| DB | 221 | LysValSerAlaAspAspValSerLeuThrGlyAlaValSerLeuAlaSerMetLeu | 240 |
| QY | 721 | GATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG | 780 |
| DB | 241 | GlnIlePheLeuLeuGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln | 260 |
| QY | 781 | GATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG | 840 |
| DB | 261 | AspSerIleGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnIlePheIleLeuLeu | 280 |
| QY | 841 | CGTACCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 900 |
| DB | 281 | ArgThrProIleValAlaAlaArgSerArgAlaIleProLeuLeuAspLeuIleLysThr | 300 |
| QY | 901 | TTGAGGAGGATCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG | 960 |
| DB | 301 | LeuThrProHisProGlnLysGlnAlaIleGlyValIleLeuProIleSerValIle | 320 |
| QY | 961 | TTTATCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG | 1020 |
| DB | 321 | PheIleAlaGlyHisAspThrAsnAlaAlaAlaAlaAlaAlaAlaAlaAlaAla | 40 |
| QY | 1021 | ACGTTTCGCGTCAGCGGATGATGATGATGATGATGATGATGATGATGATGATG | 1080 |
| DB | 341 | ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnValIleGlnAlaIle | 400 |
| QY | 1081 | CGTCGGCTAAGCGATACACAGCGAGTCTGATTCATCTTCTGCTGCTGCTGCTG | 1140 |

| | | | |
|----|------|--|------|
| QY | 781 | GATTCACAGCATGACAAACGCTTGTAAAGTTTGACAAAGCGCAATTTGATTTGCTACAA | 840 |
| DB | 261 | ASPSRHISLITTPASnThrLeuLeuSerLeuHisAsnAlaGlnPhenylLeuLeuGln | 280 |
| QY | 841 | CGACGCGACAGCTTCGCGACGCGCGACGCGCTATTAGATTGTGACAGACAGCG | 900 |
| DB | 281 | AcThrProGluValAlaLeuSerAlaThrProLeuLeuAspLeuIleLysThrAla | 300 |
| QY | 901 | TTGACGCCCAATCGACGCAAAAACAGCGGTATGCTGTGACATPAGCCACTTCAGTGGTGG | 960 |
| DB | 401 | LeuThrProHisSerProGlnLysGlnAlaTyrGlyValThrLeuProHisSerValLeu | 320 |
| QY | 961 | TTTATCGCGGACACGATCACTAATCTGCAAAATCGCGCGCGCATGAGGTTCACATCG | 1020 |
| DB | 421 | PhoLLeuAspLysHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTyr | 340 |
| QY | 1021 | ACGTTTCGCGTCACGCGGATACACACGCGGCGACGTCGTGAACGTGGTGTTTAAAGCTCG | 1080 |
| DB | 441 | ThrLeuProGlnTyrLeuTyrAspAsnThrProGlnLysGlnLeuValPheSerLeuTyr | 360 |
| QY | 1081 | CGTCGCTAAGCGATACACGCGACGTGATTCAGTTTCGCTGGTTCGACAGCTTTACAG | 1140 |
| DB | 461 | ArgLeuLeuSerAspAsnSerGlnTyrPheGlnValSerLeuValPheGlnThrLeuGln | 380 |
| QY | 1141 | CATATGGTGTATAAACGGCGGTGTTCATTAAATACGCGCGCGACAGGTGCAAACTGACG | 1200 |
| DB | 481 | GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlnLysLeuValLysLeuThr | 400 |
| QY | 1201 | CTGCGAGAGATGTAAACAGCAAAATCGAGGCGATGCTGTGTGGTCAGAGTTTACGCAA | 1260 |
| DB | 401 | LeuAlaLeuLysGlnLeuLeuLeuArgAsnAlaGlnLeuTyrMetCysSerLeuAlaGlyThrGln | 420 |
| QY | 1261 | ATGCTGAATGAGACACGATACCGCGGTGTGAGTTTG | 1296 |
| DB | 421 | IleValAsnGlnAlaGlnTyrProValCysSerLeu | 432 |

Search completed: December 13, 2002, 16:01:22
Job time : 79 secs

